



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 167420

TO: Lorraine Spector
Location: 4d55 / 4c70
Monday, October 03, 2005
Art Unit: 1647
Phone: 571-272-0893
Serial Number: 09 / 674377

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

Rest 10/4/05

This Page Blank (uspto)

167420

STIC-Biotech/ChemLib

From: Spector, Lorraine
Sent: Friday, September 30, 2005 1:42 PM
To: STIC-Biotech/ChemLib
Subject: SEARCH request for Serial No. 09/674377

STIC,

Please search SEQ ID NO: 1 and 2

-pending

-issued

-commercial

Also, please align the two sequences with each other.

Thanks.

Lorraine Spector
571-272-0893
U.S. Patent and Trademark Office
Art Unit 1647
lorraine.spector@uspto.gov
Office: Remsen 4D-55
Mailbox: Remsen 4C-70

Searcher: Jan
Searcher Phone: 22504
Date Searcher Picked up: 10/3/05
Date completed: 10/3/05
Searcher Prep Time: 10
Online Time: 20

Type of Search
NA# _____ AA# ✓
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: ✓
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspio)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:29 ; Search time 40.225 Seconds
(without alignments)
1069.209 Million cell updates/sec

Title: US-09-674-377b-1

Perfect score: 2604
Sequence: 1 ERKRRNTIHEFKSAKTLT.....IPWDYCPISRCGDTPTTIV 447

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2601	99.9	728	1 JH0579	hepatocyte growth
2	2408	92.5	728	1 A60185	hepatocyte growth
3	2401	92.2	728	1 A35644	hepatocyte growth
4	1864.5	71.6	710	1 I51283	hepatocyte growth
5	1737.5	66.7	411	2 I51285	hepatocyte growth
6	1180.5	45.3	711	1 A47136	macrophage-stimula
7	1144	43.9	716	1 JCS061	macrophage-stimula
8	1135	43.2	716	1 A40332	macrophage-stimula
9	1066	40.9	790	1 PLPG	plasma (EC 3.4.21
10	1059	40.7	812	1 PLMS	plasma (EC 3.4.21
11	1053.5	40.5	810	2 I46260	plasma (EC 3.4.21
12	1044.5	40.1	810	2 B30848	plasma (EC 3.4.21
13	1042	40.0	810	2 PLHU	plasma (EC 3.4.21
14	1038	39.9	812	1 PLBO	plasma (EC 3.4.21
15	943.5	36.2	2869	2 T18518	apolipoprotein(a)
16	879	33.8	4548	1 S00657	apolipoprotein(a) (EC
17	843.5	32.4	1420	2 A32869	apolipoprotein(a)
18	457	17.5	460	2 B61545	plasma (EC 3.4.21
19	454	17.4	455	2 A61545	plasma (EC 3.4.21
20	408.5	15.7	336	2 S33879	plasma precursor
21	357.5	13.7	625	1 TBBO	thrombin (EC 3.4.2
22	353	13.6	169	2 A40522	plasma (EC 3.4.21
23	352.5	13.5	618	2 A35827	thrombin (EC 3.4.2
24	350	13.4	622	1 TBHU	thrombin (EC 3.4.2
25	344	13.2	617	2 S10511	thrombin (EC 3.4.2
26	307.5	11.8	559	1 A35029	t-plasminogen acti
27	303	11.6	562	1 UKHUT	t-plasminogen acti
28	297.5	11.4	559	1 A29941	t-plasminogen acti
29	270	10.4	120	2 E61545	plasma (EC 3.4.21

30	268	10.3	123	2 C61545	plasma (EC 3.4.21
31	260.5	10.0	89	2 A60140	plasma (EC 3.4.21
32	235.5	9.0	946	1 A47299	ror-related recep
33	230	8.8	943	2 B45082	neurotrophic recep
34	228.5	8.8	603	2 S28941	coagulation factor
35	225.5	8.7	291	2 I38098	t-plasminogen acti
36	217	8.3	937	2 A45082	neurotrophic recep
37	204	7.8	655	1 A46688	neurotrophic recep
38	195	7.5	615	1 KFHU12	coagulation factor
39	192.5	7.4	558	2 JCS878	plasma hyaluronan-
40	189	7.3	593	2 S45281	coagulation factor
41	186.5	7.2	560	1 JCS795	plasma hyaluronan-
42	182	7.0	806	2 T18840	hypothetical prote
43	159	6.1	685	1 A48289	neurotrophic recep
44	158.5	6.1	433	1 JN0560	u-plasminogen acti
45	158	6.1	442	1 UKPG	u-plasminogen acti

ALIGNMENTS

RESULT 1

JH0579
hepatocyte growth factor precursor [validated] - human
N/Alternate names: hepatotelin A; scatter factor
C/Species: Homo sapiens (man)
C/Date: 17-Aug-1992 #sequence revision 17-Aug-1992 #ext change 09-Jul-2004
C/Accession: JH0579; J00333; A41140; B36677; A36677; A3512; A39006; PH0114; A37796; S06
R;Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A/Title: Organization of the human hepatocyte growth factor-encoding gene.
A/Reference number: JH0579; MUID:91340155; PMID:1831432
A/Accession: JH0579
A/Molecule type: DNA
A/Residues: 1-728 <SER>
A/Cross-references: UNIPROT:P14210; DDBJ:D90318
A/Note: the authors translated the codon GAA for residue 662 as Gly
R;Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
submitted to JIPID, March 1991
A/Description: Organization of the human hepatocyte growth factor-encoding gene.
A/Reference number: J00333
A/Accession: J00333
A/Molecule type: DNA
A/Residues: 1-481, RT, 484-728 <SE2>
R;Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A/Title: Evidence for the identity of human scatter factor and human hepatocyte growth f
A/Reference number: A41140; MUID:9134393; PMID:1831266
A/Accession: A41140
A/Molecule type: mRNA
A/Residues: 1-728 <WEI>
A/Cross-references: GB:M73239; NID:9337935; PIDN:AAA64239.1; PID:9337936
R;Seki, T.; Ihara, I.; Sugimura, A.; Shimomishi, M.; Nishizawa, T.; Asami, O.; Hagiya, T
Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A/Title: Isolation and expression of cDNA for different forms of hepatocyte growth fact
A/Reference number: A36677; MUID:91025062; PMID:2145836
A/Accession: B36677
A/Molecule type: mRNA
A/Residues: 1-728 <SE3>
A/Cross-references: GB:M60718; NID:9184031; PIDN:AAA52648.1; PID:9184032
A/Accession: A36677
A/Molecule type: mRNA
A/Residues: 1-161, 167-728 <SE4>
A/Cross-references: EMBL:X16323
A/Experimental source: leukocyte
R;Miyazawa, K.; Tsudouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakay
Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A/Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth fac
A/Reference number: A3512; MUID:89392017; PMID:2528952
A/Accession: A3512
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-728 <MIY>

A:Cross-references: GB:M29145; NID:g184041; PIDN:AAA52650.1; PID:g306646
R:Rubin, U.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hild
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A:Title: A broad-spectrum human lung fibroblast-derived variant of hepatoc
A:Reference number: A39006; MUID:91110540; PMID:1824873
A:Accession: A39006
A:Molecule type: mRNA
A:Residues: 1-161,167-728 <RUB>
A:Cross-references: GB:M55379
A:Experimental source: embryonic lung
R:Oshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama,
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A:Title: Identification of the N-terminal residue of the heavy chain of both native and
A:Reference number: PH0114; MUID:91207365; PMID:1826837
A:Accession: PH0114
A:Molecule type: protein
A:Residues: 32-43;53-58 <YOS>
A:Experimental source: plasma
R:Weidner, K.M.; Behrens, J.; Vandekekerhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi
A:Reference number: A37796; MUID:91035621; PMID:2146276
A:Accession: A37796
A:Molecule type: protein
A:Residues: 86-91;329-344;356-363,'XX',366-370;425-434;442-447,'X',449-450;543-546,'X',5
R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimomishi, W.; Sugimura, A.; Tash
Nature 342, 440-443, 1989
A:Title: Molecular cloning and expression of human hepatocyte growth factor.
A:Reference number: S06794; MUID:90066676; PMID:2531289
A:Accession: S06794
A:Molecule type: mRNA
A:Residues: 1-31,'HK',34-77,'N',79-292,'V',294-299,'W',301-316,'A',318-335,'K',337-386,'
A:Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA34387.1; PID:g32082
A:Experimental source: liver
A:Note: the authors translated the codon CAG for residue 727 as Glu
R:Hartmann, G.; Mäldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
A:Reference number: I59214; MUID:93087571; PMID:1280830
A:Accession: I59214
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-288,'ET','HAR>
A:Cross-references: GB:L02931; NID:g184033; PIDN:AAA52649.1; PID:g184034
R:Myazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
Eur. J. Biochem. 197, 15-22, 1991
A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor
A:Reference number: S15443; MUID:91200041; PMID:1826653
A:Accession: S15443
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288,'ET','MTY2>
A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA40802.1; PID:g32084
R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning
A:Reference number: I52253; MUID:92062058; PMID:1835583
A:Accession: I52253
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 161-166 <SHI>
A:Cross-references: GB:S62561; NID:g237996; PIDN:AAB20169.1; PID:g237997
C:Genetics:
A:Gene: GDB:HGF
A:Cross-references: GDB:127524; OMIM:142409
A:Map position: 7q21.1-7q21.1
A:Intons: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A:Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; krigle homolo

C/Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyrrrolidone

F.1-31/Domain: signal sequence #status predicted <SIG>

F.32-494,495-728/Product: hepatocyte growth factor #status experimental <MAT>

F.32-494/Domain: alpha chain #status experimental <ACH>

F.128-206/Domain: kringle homology <KR1>

F.211-288/Domain: kringle homology <KR2>

F.305-383/Domain: kringle homology <KR3>

F.391-469/Domain: kringle homology <KR4>

F.495-728/Domain: beta chain #status experimental <BCH>

F.495-716/Domain: trypsin homology <TRY>

F.32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment

F.1294,402,566,653/Binding site: carboxylic acid (Asn) (covalent) #status predicted

F.487-604/Disulfide bonds: #status predicted

Query Match	99.9%	Score 2601	DB 1	Length 728
Best Local Similarity	99.8%	Pred. No. 1.9e-176		
Matches 446	Conservative 1	Mismatches 0	Indels 0	Gaps 0

Query Match 99.9%; Score 2601; DB 1; Length 728;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTIHEKKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNKGLPEPTCKAFVFDK 600
:|||||
Db 32 QRKRNTIHEKKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNKGLPEPTCKAFVFDK 910

32 QKRRNTIHEFKKSAKTYLIKIDPALKIKTKKVNNTADQCANRCTRNKGLPFTCKAFVFDK 91

QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCOP 12

Db 92 ARKQCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 15

QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEEGGMCFSTNPVRYEVC DIPQSEVEC 18

D6 152 WSSMIPHEHSFLPSSYRGRDLQENYCRNPRGEGGFWCFISNPEVKYEVCDIPQCSVEVC 21

181 M I C N G E S I R G L M D H I E S G K L C Q R W D H Q I P H K H A T D P E K I P D V A G F D D N I C A N F D G Y F R F W C 23

212 MICROSERIES IN ECONOMIC RESEARCH: A REVIEW OF THE LITERATURE

241 I I D D F H I N M E I C H A I N I C H A D N I F N D I D V E L E I E C I Q Q Q Q Q Q I N O I V N I I M N O I C Q R M D Q

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
[
\
]
^
_
`
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
{
|
}
~

453 OVERCUT TUNING DIODES 476

RESULT 2
A60185

N: Alternate names: hepatocellular carcinoma

C:\species: mus musculus (mouse mouse)
C:\Date: 03-Mar-1993 #sequence revision 26-May-1994 #text change 09-Jul-2004

R;Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.

A;Title: Identification of mouse mammary fibroblast-derived mammary growth factors

A;Accession: JC2117

A;Residues: 1-728 <SAS2>

A; Experimental source: fibroblast, COS-1 cell

A;Accession: PC2064

A;Residues: 496-504 <SA2>

Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990

```

A>Title: Purified scatter factor stimulates epithelial and vascular endothelial cell mig
A:Reference number: A60185; MUID:90377927; PMID:2144630
A:Accession: A60185
A:Molecule type: protein
A.Residues: 'X','184-188','KX','191-192','X','194','XX','197,357-364','XX','367,375-377','E','379','
R.Liu, Y.; Michalopoulos, G.K.; Zaneagar, R.
Biochim. Biophys. Acta 1216, 299-303, 1993
A>Title: Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth
A:Reference number: S43416; MUID:94060105; PMID:8241272
A:Accession: S43416
A>Status: preliminary
A:Molecule type: mRNA
A.Residues: 1-728 <LIU>
A:Cross-references: EMBL:X72307
R.Liu, Y.
submitted to the EMBL Data Library, May 1993
A:Reference number: S45521
A:Accession: S45521
A>Status: preliminary
A:Molecule type: mRNA
A.Residues: 1-563,'H','565-728 <LI2>
A:Cross-references: EMBL:X72307
R.Coffer, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
Biochem. J. 278, 35-41, 1991
A>Title: Purification and characterization of biologically active scatter factor from r
A:Reference number: S17173; MUID:91354223; PMID:1831975
A:Accession: S17173
A:Molecule type: protein
A.Residues: 496-517,'T','519 <COF>
R.Gherardi, E.; Stoker, M.
Nature 346, 228, 1990
A>Title: Hepatocytes and scatter factor.
A:Reference number: S10966; MUID:90326152; PMID:2142751
A:Accession: S10966
A>Status: preliminary
A:Molecule type: protein
A.Residues: 496-507,'X','509-512','I','514-516','X','518-519 <NAT>
R.Plaschke-Schluter, A.; Behrens, J.; Gherardi, E.; Bitchmeier, W.
J.Biol. Chem. 270, 830-836, 1995
A>Title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
A:Reference number: I48758; MUID:95122532; PMID:7822318
A:Accession: I48758
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A.Residues: 1-30 <RES>
A:Cross-references: EMBL:X81630; MID:9673451; PIDN:CA57286.1; PID:9673452
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:function:
A:description: stimulates mitosis of hepatocytes and other cells
A>Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyn
F.1-35/Domains: signal sequence #status predicted <SIG>
F.56-95,496-728/Product: hepatocyte growth factor #status predicted <MAT>
F.56-95/Domains: hepatocyte growth factor alpha chain #status predicted <ACH>
F.129-207/Domains: kringle homology <KR1>
F.212-289/Domains: kringle homology <KR2>
F.306-384/Domains: kringle homology <KR3>
F.392-470/Domains: kringle homology <KR4>
F.496-728/Domains: hepatocyte growth factor beta chain #status predicted <BCH>
F.496-719/Domains: trypsin homology <TRY>
F.496-Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F.295_403_569_656/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.488-607/Disulfide bonds: #status predicted

Query Match          92.5%; Score 2408; DB 1; Length 728;
Best Local Similarity 91.1%; Pred.No. 8.6e-163;
Matches 407; Conservative 21; Mismatches 19; Indels 0; Gaps 0;

QY      1 ERRRRNTIHFFKSAAKTLTKRIDPALKTITKKVNVNADOCARCTNNKGIPFTCKAPVRPK 60
DB      33 QKRRRTTHFFKSAKTLTKIEDPLTKIKTKVNSADSCANNCRINRGFTTCAPAFPRK 92

```

QY	61	ARXCIWPPFMSMSGKYKKEFGHEFDLYENNDYIRNCLIGKRGSEYKGVSTKSGICOP	120
Db	93	SRKCCYWPFPMSMSGKVGKGFGEHFDLYENNDYIRNCLIGKRGSEYKGVSTKSGIKCOP	152
QY	121	WSSMI PHEHSPFLPSSYSGKDLQENYCRNPGEESGPMCFTSNPPVRYEVCDIPOCSEVEC	180
Db	153	MNSMI PHEHSPFLPSSYSGKDLQENYCRNPGEESGPMCFTSNPPVRYEVCDIPOCSEVEC	212
QY	181	MTNGESYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDDNYCRNPDQPPWC	240
Db	213	MTNGESYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDDNYCRNPDQPPWC	272
QY	241	YTLBPHTPWECAIKTCAADNTMNDTVPLEHTETCIGQSGGSEYRGTNTNMGICQORWDS	300
Db	273	YTLBPHTPWECAIKTCHSAVNNETDVPMEHTETCIGQSGGSEYRGTNTNMGICQORWDS	332
QY	301	QYPHEHMDTPENFKCKDLRENYCRNPDGSESPMCFITDPNIRVYCSQIIPNCDMSHQDC	360
Db	333	QYPKHHDTPENFKCKDLRENYCRNPDGABSPMCFITDPNIRVYCSQIIPNCDSSQDC	392
QY	361	YRGNKKNYMGNLQSTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNEYCRNPDDAHGPW	420
Db	393	YRGNKKNYMGNLKSTRSGLTCSMWDKNMEDLHRHIFWEPDASKLKNKYCRNPDDAHGPW	452
QY	421	CYTGNPLIPMDYCPISRCEGDTTPTIV	447
Db	453	CYTGNPLIPMDYCPISRCEGDTTPTIV	479

RESULT 3

hepatocyte growth factor precursor - rat
 N.Alternate names: hepatopietrin A; scatter factor
 C.Species: Rattus norvegicus (Norway rat)
 C.Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 09-Jul-2004
 C.Accession: A35644; S13211
 R.Prachin, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimomishi, M.; Shimizu, S.; Nakamura
 P.Taochi, R. Acad. Sci. U.S.A. 87, 3200-3204, 1990
 A.Title: Deduced primary structure of rat hepatocyte growth factor and expression of the
 A.Reference number: A35644; MUID:90222197; PMID:2139229
 A.Accession: A35644
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-728 <TMS>
 A.Cross-references: UNIPROT:P17945; GB:D90102; GB:M32987; NID:G220766; PIDN:BA14133.1;
 A.Note: the authors translated the codon GAG for residue 70 as Gln, GAC for residue 417
 R.Okajima, A.; Miyazawa, K.; Kitamura, N.
 Eur. J. Biochem. 193, 375-381, 1990
 A.Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA du
 A.Reference number: S13211; MUID:91031482; PMID:2146117
 A.Accession: S13211
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-728 <OK>
 A.Cross-references: EMBL:X54400; NID:956353; PIDN:CA18266.1; PID:94539554
 C.Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C.Function:
 A.Description: stimulates mitosis of hepatocytes and other cells
 A.Note: does not have proteinase activity
 C.Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homol
 C.Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; py
 F.1-132/Domain: signal sequence #status predicted <SIG>
 F.56-495/Product: hepatocyte growth factor #status predicted <MAT>
 F.56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F.129-207/Domain: kringle homology <KR1>
 F.212-289/Domain: kringle homology <KR2>
 F.306-384/Domain: kringle homology <KR3>
 F.392-470/Domain: kringle homology <KR4>
 F.496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F.496-719/Domain: tyrosin homology <TRY>
 F.333/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte
 F.1295/403,569,656/Binding site: carbonylated (Asn) (covalent) #status predicted
 F.1488-607/Disulfide Bonds: #status predicted

```

Query Match      92.2%; Score 2401; DB 1; Length 728;
Best Local Similarity 90.4%; Pred. No. 2.7e-162;
Matches 404; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY      1 ERKRNTTHERKKSAKTLIKIDPALKTKTKKNTTAOQANRCRNRNGLEPTCAAPFDK 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      33 QKRRNTTHERKSKSKTLTKEDPLVKTKTKVNSABECARCRIRNGPPTCAAPFDK 92

QY      61 ARKOCIMPEPFMSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKRGSKGTVSITKSGICOP 120
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      93 SRKRCYWFPPFMSMSGVAKGFGHEFDLYENKDYIRNCIIIGKSGSKGTVSITSGIKCP 152

QY      121 WSMITPHEHSPFPSSRYGKDIQENYCRNPRGEGGPMCFITSPNRYVEVCDIPQCSVEEC 180
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      153 WMSMTPEHSEFLPSSRYGKDIQENYCRNPRGEGGPMCFITSPNRYVEVCDIPQCSVEEC 212

QY      181 MTCNGESRGIMDPHESKICORDMDHOFPHHKFLPEHYPRPKGPDMDYCRNPDGAPRPMC 240
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      213 MTCNGESRGIMDPHESKTCQDMDQOTPHHKFLPEHYPRPKGPDMDYCRNPDGAPRPMC 272

QY      241 YTLPPHTFWEYCAIKTTCADNTMNDVDVPLETTECTIOGGEYRGTVNTIWMGILPCQRWDS 300
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      273 YTLDPDTFWEYCAIKMGCHSAVNETDVPMETTECIKKGGEYRGTVNTIWMGILPCQRWDS 332

QY      301 QYPHEHDMTPENFKCKDIRENYCRNPDGSESPWCFETTDENIRVYGCSQIPNCDSHGQDC 360
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      333 QYPHGHDTTPENFKCKDIRENYCRNPDGSESPWCFETTDENIRVYGCSQIPKCDVSSGQDC 392

QY      361 YRGNKANVMGNLSQTRSGITCSMDMDXNMEDLHRHIFWEPDASXJLNENYCRNPDGDAHGPW 420
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      393 YRGNKANVMGNLSKTRSGITCSMDMDXNMEDLHRHIFWEPDASXJLTNYCRNPDDAHGPW 452

QY      421 CYTGNPLIIPMDYCPISRCGGDTTPTIV 447
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      453 CYTGNPLIPMDYCPISRCGGDTTPTIV 479

RESULT 4
151283
hepatocyte growth factor precursor - clawed frog
N.Alternate names: hepatoinetin A; scatter factor
C.Species: Xenopus sp. (clawed frog)
C.Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-May-2004
C.Accession: 151283
R.Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiokawa, K.
Mech. Dev. 49, 123-131, 1995
A.Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ear
A.Reference number: 151283; MUID:95567690; PMID:7748783
A.Accession: 151283
A>Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-710 <NAX>
A.Cross-references: GB:S77422; NID:9989932; PIDN:AB34354.1; PID:9989933
A.Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleotide
C.Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C.Function:
A.Description: stimulates mitosis of hepatocytes and other cells
A.Note: does not have proteinase activity
C.Superfamily: hepatocyte growth factor/mecrophage stimulating protein 1; kringle homoloid
C.Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle
F:42-477, 478-709/Product: hepatocyte growth factor #status predicted <MNT>
F:442-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:115-193/Domain: kringle homology <KR1>
F:198-275/Domain: kringle homology <KR2>
F:289-367/Domain: kringle homology <KR3>
F:375-453/Domain: kringle homology <KR4>
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F:478-700/Domain: trypsin homology <TRY>
F:522,128,281,322,319,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status pre
F:470-588/Disulfide bonds: #status predicted

```

Best Local Similarity 69.3%; Pred. No.2,3e-124; Matches 305; Conservative 64; Mismatches 68; Indels 3; Gaps 1;

```
Oy      KRNNIHEFKSAKATTLIKIDPALKITKKVNTADOCANRCTRNKGLEPTCKAFVPDKAR   62
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     21 KKRNFADFYYKAETTLRLINKALEVKTKMFNTEHCARKCSNRNGLPFTCKAFAFDKI    80
Oy      KOCLMFPNPMSSGVYKEFGFEBDLYENKOYLINCIIIGKRSYKTGYITKSQIKCOBWS   122
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     81 KRCHWFSFNWMAGIKDKXDIISFDLXEKKDYIRDCIHGKSNRYGTENVTRKGLACOPWN   140
Oy      SMIPEHSFLPSYGKDLOEYNCRNPGBEGGPWCFTSPREVEYCDDIPCESEVCMT     182
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     141 SMIPHESFLPYTYGKDLKENYCRNPKEBGSGPWCFTSPEVRHHVDCDILPCGEVDCT    200
Oy      NCGESYRGIMDTSEGSKIQQRWHDQTPHRHKLPERYPDPGFNDNCRNPDGQPBMCYT   242
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     201 CNGEHYRGMDVTSEGSKEQQRDLORPHGHKFREFEPYPNKGDMDNVCRNPDGKSRPMCYT    260
Oy      LDPTHREWCAIKTCADNTMANDTDVPLETTCCIQGGEGYRGIVNTIWNGIPCORMDSQY   302
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     261 LDPDTSWECAIKCPVHSVINVNTDI---TWCKCGKGGEGYRGVSITYNGICCCRMSDF     317
Oy      PNEHMTPENPKCXLRENKYCRNPDESSEPWCCTTPNIINVGYSOIPLNDMSGHODCYR   362
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     318 PHLNHFTEPENYCKDISENYCRNPDESSEPWCCTTPNIRIGHOSQLKCKCASMQOEcy    377
Oy      NGGRKVYMUGLSOTRSLGCISMWDKNMEDLRHIIFWPEDASKLENYCNRPDDDAHGPPCWY 422
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     378 NGSSTYKGLSTRPRLLPCSMMEKNOIDLKRHTFNEBDVISILOKNYCRNPNDNAHGPPCWY 437
Oy      TGNPLIPMDYCPISRCEGDT 442
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     438 TDDEPVMPDWICYPISRCEGDT 457
```

RESULT 5
151285

hepatocyte growth factor/scatter factor - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #ext_change 09-Jul-2004
C:Accession: FJ151285
R:Straitt, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Cherard Development 121, 813-824, 1995
A>Title: A role for HGF/SF in neural induction and its expression in Henzen's node during A.Reference number: FJ151285, PMID:7720585
A:Accession: FJ151285
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-411 <STR>
A:Cross-references: UNIPROT:Q90978; GB:S77480; NIH:g998675; PID:g998676
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1, kringle homology Fj124-197/Domains: kringle homology <KR2>
Fj202-279/Domains: kringle homology <KR3>
Fj296-374/Domains: kringle homology <KR3>

Query Match 66.7%; Score 1737.5; DB 2; Length 411;
Best Local Similarity 74.9%; Pred. No. 1.3e-115;
Matches 289; Conservative 47; Mismatches 45; Indels 5; Gaps 1;

```
Oy      KRRNTIEHKSAKTLLIKIDPAIKITRKVNTFADQCARNCTRNKGLEPTCKAFVPDKAR   62
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     30 KORNLHYDKYTGELMLIKVNMKTLEVTKLLNTTEQCACKSCSRNGKSFCTCKAFAYDRVT    89
Oy      KOCLMFPNPMSSGVYKEFGFEBDLYENKOYLINCIIIGKRSYKTGYITKSQIKCOBWS   122
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     90 KRCHWLIFSFLTNGVARKKDHAPDLFEKDYVNRCCIIGKAEVYGTISTISYGIQOQAOWN    149
Oy      SMIPHSHPLPSYRGKDLOEYNCRNPGBEGGPWCFTSPREVEYEVCDDIPQCESEVCMT   182
       |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     150 SMIPHEH-----SYRGKDLRENKC RNPRGEGGPWCCTTSPQMHAHVCDILPCSVECMT    204
Oy      NCGESYRGIMDTSEGSKIQQRWHDQTPHRHKLPERYPDPGFNDNYCRNPDQGPRPWCVT   242
```

```

Db      205  CNESEYRGPMDDHTESGECCRMWDLQRPBKAKFRPERPDGKFDNDNCRNPDGKLRWCY 264
      243  LDPTWMEYCAITKCADNTMNDTDVLEETECIOGOGEGYRGIVNTIMNGIPCORMDSOY 302
      265  LDPTWMEFCAITTCOVGILNSTEVAETTCIOGOGEGYRGIVNTIMSGICORMDSOY 324
      303  PHEHMTPEMFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSMGHGDQCYR 362
      325  PHQHNITPEMFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPKCDVSNEDQCYR 384
      363  GNGKNTWGNLSQTRSGITCSMDKDM 388
      385  GNGKSYMGNLSKTRFGLTCSMDKDM 410

RESULT 6
A47136
macrophage-stimulating protein 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 09-Jul-2004
C:Accession: A40331; B40331; A47136; A61395
R:Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A:Title: Characterization of the DNFI5S2 locus on human chromosome 3: identification of
A:Reference number: A40331; MUID:92002016; PMID:1655021
A:Accession: A40331
A:Molecule type: DNA
A:Residues: 1-711 <HA1>
A:Cross-references: UNIPROT:P26927; GB:M74179
A:Accession: B40331
A:Molecule type: mRNA
A:Residues: 1-711 <HA2>
A:Cross-references: GB:M74178; NID:q183976; PIDN:AAA50165.1; PID:q183977
R:Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A:Title: Cloning, sequencing, and expression of human macrophage stimulating protein (MS
A:Reference number: A47136; MUID:93340141; PMID:8393443
A:Accession: A47136
A:Molecule type: mRNA
A:Residues: 1-12, 'C', 'A'-622, 'R', '624-711' <YOS>
A:Cross-references: GB:L11924; NID:q398037; PIDN:AAA59872.1; PID:q398038
R:Skjel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 173, 1227-1234, 1991
A:Title: Macrophage stimulating protein: purification, partial amino acid sequence, and
A:Reference number: A61395; MUID:91217635; PMID:1627141
A:Accession: A61395
A:Molecule type: protein
A:Residues: 230-247,288-291, 'E', '293-295', 'X', '297-301', 'X', '303', 'E', '305', 'EX', '308-310',326-331
A:Experimental source: plasma
C:Genetics:
A:Gene: GDB:MST1, D3F15S2, DNFI5S2; HGFL
A:Cross-references: GDB:126833; OMIM:142408
A:Map position: 3p21-3p21.3
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
C:Keywords: duplication; glycoprotein; growth factor; kringle; plasma
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-483,484-711/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F:19-483/Domain: alpha chain #status predicted <ACH>
F:110-186/Domain: kringle homology <KR1>
F:191-268/Domain: kringle homology <KR2>
F:283-361/Domain: kringle homology <KR3>
F:370-448/Domain: kringle homology <KR4>
F:484-711/Domain: beta chain #status predicted <BCH>
F:484-704/Domain: trypsin homology <TRY>
F:56-78,60-66,110-166,131-169,157-181,191-268,212-251,240-263,283-361,304-343,332-355,37
F:12,296,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      45.3%; Score 1180.5; DB 1; Length 711;
Best Local Similarity 47.4%; Pred. No. 5.6e-76;
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;
```

```

      33  VNTADQANRCRRNKGLPFTCAFAVDFKARKOCLWPFNMSGSGVKEFGHEFDLYEND 92
      50  VDAEBCARCC-----GPLMDCAAFHNVNSSHCOQLLPWQHSHPTLRASGRCDLFOKDD 105
      93  YTRNCTIGKGRSGYKGTSTITKSGIKQCPWSSMIPIHEHSLPSSYRGKDLQENYCRNPRGE 152
      106  YVTCIMNNGVYRGTMATTVGGLPCQAWSHKFPNDHKYTPLRNNG--LEENFCRNPDGD 163
      153  EGGPWCFTSNPFRYAVCDIPQCEVEGCTNGESYRGIMHDTESGKICORMDHQPFRH 212
      164  PGFPWCYTTDPVAFRFQSCGKSCREACVWNGEETRGAVDTESEGCQMDLQHPHQ 223
      213  KFLPERYPDKGFPDNYCRNPDGQPRWCYTLDPHTWMEYCAITKCADNTMNDTDVLETT 272
      224  PFEFGKFLDQGLDDNYCRNPDGSESPWCYTTDPQIEREFCDLPRGSEAQPRQEA--TTV 281
      273  ECIQOGEGYRGIVNTIMNGIPCORMDSOYPHEHMTPEMFKCKDLRENYCRNPDGSESP 332
      282  SCFRKGEGYRGIVNTIMNGIPCORMDAIPHQHFTPEKFAKCKDLRENYCRNPDGSEAP 341
      333  WCFTTDPNIRVGYCSQIPNC-DMSGHGDQCYRNGKNYGNLSQTRSGITCSMDKDMEDL 391
      342  WCFTLRPGKRAAFCTQIRRCTDVRPDQCYHGAEGYRGTVSKTRKGVCQGRWASATPRK 401
      392  HRHIEWEPDASKLINENYCRNPDGDAHGPWCYGNPLIPMDYCPISRCBGDTTPPTIV 447
      402  PGFTFSEBPAQLEENFCNPNPDGSHGRCYIMDPRTPTDYCALRCADQPPSIL 457

RESULT 7
JC5061
macrophage-stimulating protein 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C:Accession: JC5061
R:Ohshiro, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, K.
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A:Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
A:Reference number: JC5061; MUID:97011126; PMID:8858136
A:Accession: JC5061
A:Molecule type: mRNA
A:Residues: 1-716 <ONS>
A:Cross-references: UNIPROT:P70521; EMBL:X95096; NID:q1669718; PIDN:CAA64473.1; PID:q1669718
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
C:Keywords: duplication; glycoprotein; growth factor; kringle
F:1-31/Domain: signal sequence #status predicted <Sig>
F:32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F:110-186/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F:191-268/Domain: kringle homology <KR1>
F:191-268/Domain: kringle homology <KR2>
F:292-370/Domain: kringle homology <KR3>
F:379-457/Domain: kringle homology <KR4>
F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F:489-709/Domain: trypsin homology <TRY>
F:12,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      43.9%; Score 1144; DB 1; Length 716;
Best Local Similarity 46.1%; Pred. No. 2.1e-73;
Matches 196; Conservative 70; Mismatches 141; Indels 18; Gaps 6;
```

```
QY 213 KFLPERYPDKGFDDNYCRNPDGQPRPWCYTLDPHTRMEYCAIKTCADN-----TMNDT 265
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 224 PFHPEKFPDKALNDYCNRPDASERPMWCYTDPNVERECDLPSGPNLPPTTKSKSQ 283
QY 266 DVPLETTECTIGOGEGRGYGTNTVINGIPQGRMSQYRPHEDMTPEPKCKDLRENYCRN 325
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 284 RNKVASNCFRGKGGEDYRGTTNTTSAGVPCQRMDAQNPQHFRFVPEKYACKDLRENYCRN 343
QY 326 PDGSEPMCFETDPPNIRVGYCSQIPNC-DMSHGQDCYRGNGKNYMGNLSTQTSGLTCSMW 384
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 344 PDGSEPMCFETSRPRGRVAFCTQIPRCTEEVVPBGCTHSGSGOYRGSVSKTRKGVQCCOH 403
QY 385 DKNMEDLHRHIFWEPDAS--SKLNENYCRNPDDAHGMPCYTGPNLIPMDYCPISRCEGDT 442
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 404 --SSETPHKKPQFTPTSPAPAGLEAFNFCRNPDDGSHGPMWCYTLDPELTFYCALKKCCDDQ 461
QY 443 TPPTIV 447
      |||:
Db 462 PPSIL 466
```

RESULT 8

```
A40332
macrophage-stimulating protein 1 precursor - mouse
N:Alternate names: hepatocyte growth factor-like protein
N:Contains: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C:Accession: A40332; #0332
R:Degeen, S.U.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
Biochemistry 30, 9781-9791, 1991
A:Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth fact
A:Reference number: A40332; MUID:92002017; PMID:1832957
A:Accession: A40332
A:Molecule type: DNA
A:Residues: 1-716 <DEG>
A:Cross-references: UNIPROT:P26928; GB:M74180; NID:g193831; PIDN:AAA50166.1; PID:g193832
A:Accession: B40332
A:Molecule type: mRNA
A:Residues: 1-18,'P',20-716 <DEG>
A:Cross-references: GB:M74181; NID:g193833; PIDN:AAA50167.1; PID:g193834
C:Genetics:
A:introns: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/1; 412/2; 458/1; 47
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
C:Keywords: duplication; glycoprotein; growth factor; kringle
F:1-31/Domain: signal sequence #status predicted <SIG>
F:19-488,489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>
F:19-483/Domain: alpha chain #status experimental <ACH>
F:110-186/Domain: kringle homology <KR1>
F:191-268/Domain: kringle homology <KR2>
F:292-370/Domain: kringle homology <KR3>
F:379-457/Domain: kringle homology <KR4>
F:484-711/Domain: beta chain #status experimental <BCH>
F:489-709/Domain: trypsin homology <TRY>
F:172,173,305,620/Binding site: carbohydrate (asn) (covalent) #status predicted
```

```
Query Match 43.2%; Score 1125; DB 1; Length 716;
Best Local Similarity 45.9%; Pred. No. 4,7e-72;
Matches 195; Conservative 69; Mismatches 143; Indels 18; Gaps 6;
QY 33 VNTADOCANRCTENKGLPTCKAFVVDKARKQCLMPFPNMSGSGVKKERGFHEFDLYEND 92
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 50 VADAECCARRC-----GRLDCAFAHNMSSHCQLLPWQHSLHTQLYHSSLCILFKKDD 105
QY 93 YVRNCTIGKRGSKYGVSTKSGIKQOPWSSMTPEHSHLPSSSYRGKDLQENYCRNPKE 152
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 YVRCTCMDGVSYRGVARTAGGLPCQAMSRFPNDHKXTPPKNG--LEENFCRNPDD 163
QY 153 EGGPMCFSTNPEYRVCEIPQSEVFCMTNGESYRGIMDHESKICIQRMWHQPHRH 212
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 164 PRGPMCYTTNRSYVFOSCGIKTCREAVCYLCSNEDYRGVDTESGRBCQRMDLQPHSH 223
```

```
QY 213 KFLPERYPDKGFDDNYCRNPDGQPRPWCYTLDPHTRMEYCAIKTCADN-----TMNDT 265
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 224 PFQPEKFLDKLNDYCNRPDASERPMWCYTDPNVERECDLPSGPNLPPTTKSKSQ 283
QY 266 DVPLETTECTIGOGEGRGYGTNTVINGIPQGRMSQYRPHEDMTPEPKCKDLRENYCRN 325
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 284 RNKVASNCFRGKGGEDYRGTTNTTSAGVPCQRMDAQNPQHFRFVPEKYACKDLRENYCRN 343
QY 326 PDGSEPMCFETDPPNIRVGYCSQIPNC-DMSHGQDCYRGNGKNYMGNLSTQTSGLTCSMW 384
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 344 PDGSEPMCFETSRPRGRVAFCTQIPRCTEEVVPBGCTHSGSGOYRGSVSKTRKGVQCCOH 403
QY 385 DKNMEDLHRHIFWEPDAS--LNENYCRNPDDAHGMPCYTGPNLIPMDYCPISRCEGDT 442
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 404 --SSETPHKKPQFTPTSPAPAGLEAFNFCRNPDDGSHGPMWCYTLDPELTFYCALORCDDQ 461
QY 443 TPPTIV 447
      |||:
Db 462 PPSIL 466
```

RESULT 9

```
PIPG
plasmin (EC 3.4.21.7) precursor - pig (fragment)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Sep-1990 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: S03733; S03737; A25834
R:Schaller, J.; Marti, T.; Roesseler, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the ca
A:Reference number: S03733
A:Accession: S03733
A:Molecule type: protein
A:Residues: 1-560 <SCH>
A:Cross-references: UNIPROT:P06667
R:Brundishol, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03737
A:Molecule type: protein
A:Residues: 1-57 <BRU>
R:Marti, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A:Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.
A:Reference number: A25834; MUID:85203907; PMID:3846533
A:Accession: A25834
A:Molecule type: protein
A:Residues: 450-790 <MAR>
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
as the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine prote
F:1-790/Product: plasminogen #status predicted <PRO>
F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>
F:1-77/Domain: activation peptide #status predicted <AP>
F:78-560/Product: plasmin chain A #status predicted <ACH>
F:84-162/Domain: kringle homology <KR1>
F:166-243/Domain: kringle homology <KR2>
F:256-333/Domain: kringle homology <KR3>
F:358-435/Domain: kringle homology <KR4>
F:450-790/Product: miniplasminogen #status experimental <MIN>
F:461-540/Domain: kringle homology <KR5>
F:561-790/Product: plasmin chain B #status experimental <BCH>
F:561-783/Domain: trypsin homology <TRY>
F:30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,305
bonds: #status predicted
F:602,645,740/Active site: His, Asp, Ser #status predicted
```


	Query Match	40.9%	Score 1066;	DB 1:	Length 790;
	Best Local Similarity	44.6%;	Pred. No.	7.8e-68;	
	Matches	195;	Conservative	65;	Mismatches 137; Indels 40; Gaps 13;
Oy	ALTKIKKKVNTAOCANRCRNKGLPFTCKAFVDPKARKOCLMFPFSSMSGVKKEFGHE	84			
Dd	SLSRKQVAARSVECAAKCEAEHTN--FLICRAFYHSHXDOQCVMAMENKTSPIAM--RD	71			
Oy	FDLYENKDVIIRNCITGKGSRYSKGTVS.ITTSGIHKQPWSMT.PHEHSFILPSRYGDLOEN	144			
Dd	VVLFEKRILYLSECKTGNGKNVRGTTSTKTSKVIGIQKSIVSPPHI.PKYAPEPEPLAGLEEN	131			
Oy	YCNRNPGEGGGPCFTSNPEVRVYECDIPOSSEVCMTCNESRGLMDHTESSGIQRW	204			
Dd	YCNPNPENDEGPWCYTTPDPTRFDCOIPEC-EEDCMECSGEHYEGKISKTMWG.IECSOW	190			
Oy	DHOTPRRHKLPERYPDPKCFDDNYCRNDPGOPRMWCYTLLDPTHREYAICKANTMYMD	264			
Dd	GGSPPAHGHLPBKPFRPNLKMAYGRMPDGEPWRMCFTTDNKRRKFCDIPRC--TTPRP	248			
Oy	TUVPLETECTICOGQEGRGTVNTIANGIPCQRWDSDYPHSHDMTPENFKCKDLKENYCR	324			
Dd	TSCGP--TYOLCLKRGENVRTGVSVTAGHTCQRMSAQSPHHGNRPENPFPCNLIEYNCR	306			
Oy	NPDGSSPMFCFTDPDNTRVGCGSIENCDS-----H-----GQDYCHNGKN	367			
Dd	NBDGETAPMCYYTIDSEVRWDYC-KIPSGSSTTSTEHLDAVPBPQTVAODCCYGNGNS	365			
Oy	YMGNLSQTSQSGLTCSMWKDKMEDLRHIFWE-----PDASKLNENYCRNPDDAHGPC	421			
Dd	YGYTSTTTITGRKCOSNVSWTP-HRH---EKTEGNFNAC-LTNMYCRNPAD-KSPWC	418			
Oy	422 YTGNNPLIPMDYCPISRC	438			
Dd	419 YTTPDRVRWEYCNLKCC	435			
	RESULT 10				
	PLMS				
	Plasmin (EC 3.4.21.7) precursor - mouse				
	N:Contains: angiotensin; plasminogen				
	C:Species: Mus musculus (house mouse)				
	C>Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004				
	C:Citation: A38514; S48202; S48203				
	R:Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.				
	Genomics 8, 49-61, 1990				
	A>Title: Characterization of the cDNA coding for mouse plasminogen and localization of t				
	A:Reference number: A38514; MUID:91184812; PMID:2081600				
	A:Accession: A38514				
	A:Molecule type: mRNA				
	A:Residues: 1-812 <DBG>				
	A:Cross-references: UNIPROT:P20918; GB:J04766; NID:S200402; PIDN:AAA50168.1; PTD:g200403				
	R:Lilien, H.R.; van Hoef, B.; Beelen, V.; Collen, D.				
	Eur. J. Biochem. 224, 863-871, 1994				
	A>Title: Characterization of the murine plasma fibrinolytic system.				
	A:Reference number: S48202; MUID:95010076; PMID:7523120				
	A:Accession: S48202				
	A:Molecule type: protein				
	A:Residues: 20-25<LIU>				
	A:Accession: S48203				
	A:Molecule type: protein				
	A:Residues: 22-27>				
	C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many o				
	C:Comment: Plasminogen is converted into plasmin by plasminogen activators, both plasm				
	e inhibitor, the activation involves also removal of the activation peptide.				
	C:Comment: Streptelysin I (see PI-RKCMSSI) acts on plasminogen to produce angiotatin. To				
	efful in treating solid tumors.				
	C:Function:				
	A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a v				
	n the walls of the graafian follicle; also activates the urokinase-type plasminogen act				
	A:Pathway: fibrinolysis				
	C:Superfamily: plasmin; kringlike homology; plasminogen-related protein precursor homology				

C:Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyd

F:1-36/Domain: plasminogen-related protein precursor homology <PLBP>

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-812/Product: plasminogen #status predicted <PRO>

F:20-96/Domain: activation peptide #status predicted <APT>

F:79-466/Product: angiotatin #status predicted <AST>

F:97-581,582-812/Product: plasmin #status predicted <MAT>

F:97-581/Domain: chain A #status predicted <ACH>

F:103-181/Domain: kring1e homology <KR1>

F:185-262/Domain: kring2 homology <KR2>

F:275-352/Domain: kring3 homology <KR3>

F:377-454/Domain: kring4 homology <KR4>

F:481-556/Domain: kring5 homology <KR5>

F:582-812/Domain: chain B #status predicted <BCH>

F:582-805/Domain: trypsin homology <TRY>

F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32

bonds: #status predicted

F:78-79/Cleavage site: Glu-Asn (astromelysin 1) #status predicted

F:136,309/Cleavage site: carbohydrate (asn) (covalent) #status predicted

F:466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted

F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental

F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 40.7%; Score 1059; DB 1; Length 812;

Best Local Similarity 43.3%; Pred. No. 2.5e-67; Mismatches 145; Indels 38; Gaps 12;

Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

30 TKKVNTA--DOCANRCTRNKGLPFTCKAFVFPKARQCLFPPNMSGGYKKEFGHEFD 86

37 TKKQLAGVSDCLAKCEB--TDVFCRSFQYHSKEQCVIAMNSKTSIIIM--RDVI 92

87 IYENKDYIRNCITIGKRSYKGVSTIKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146

93 LFEKRVYLSECKTGIGNGRGRTWSRTKSGVACQKMGATFPHVPYVPSSTHNEGLEENYC 152

147 RNPREEGPGMCKFTSNPEYREYVCDIPQCEVEVCMTCNGESYRGLMDHTESGKICQRDH 206

153 RNPNDDEQGPWCYTTDPDKRYDCNIPCEBE--RCMYSGEGYEKISKTMGLDQAWDS 211

207 QTPHRRHKFLPERYPDPKGFDDNYCRNPDPGQRPWCYTTDLPTTRMEYCAIKTCADNTMNDTD 266

212 QSPHAGYIPARFSPSKNLKMNVCNHPDGEPRPFCFTTDPTRKWEYCDIPRCT---TPPP 267

267 VPLETTECIQGGEGYRGTVNTIANGIPCORWDSQYVHEHDMTPENFKCKDLRENYCNP 326

268 PPSPTQYCLKGGKGENYRGTVSVTSVSGTKQRMSEQTPHRRNRPENFPCKNLENYCNP 327

327 DGSESPWCTTTPNIRVYGCQSIQPNCDMSHG-----QDCYRGKGYM 369

328 DGETAPWCYTTDSQLMWEYC--ELPSCSSASPDQSDSSVPEEQTPVVOECYQSDGOSYR 386

370 GNLSQTRSGLTGCSMDKNNMEDLRHIFWE--PDASKLNEYCNPDDDAGPWCYGNP 426

387 GTSSTTTITTKKQSQMAAMP--HRHSKTPENFPDAG--LENNYCNPNPDG--KGPWCYTTDP 442

427 LIPMDYCPISRCGDTTPPIV 447

443 SVRMETCNLKRC--SETGSGSV 462

RESULT 11

146260

plasmin (EC 3.4.21.7) precursor - western European hedgehog

C:Species: Erinaceus europaeus (western European hedgehog)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004

C:Accession: I46260

R:Lawm, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; For

J. Biol. Chem. 270, 24004-24009, 1995

A:Title: The recurring evolution of lp(a): Insights from cloning of hedgehog apolipopro

A:Reference number: I46259; MUID:96025778; PMID:7592357

A:Accession: I46260

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-810 <LMA>
 A:Cross-references: UNIPROT:Q29485; EMBL:U33171; NID:G1046360; PID:G1046361
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C:Keywords: hydrolase; serine proteinase
 F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
 F:103-181/Domain: kringle homology <KR1>
 F:185-262/Domain: kringle homology <KR2>
 F:275-352/Domain: kringle homology <KR3>
 F:379-456/Domain: kringle homology <KR4>
 F:482-561/Domain: kringle homology <KR5>
 F:582-803/Domain: trypsin homology <TRY>

Query Match 40.5%; Score 1053.5; DB 2; Length 810;
 Best Local Similarity 44.3%; Pred. No. 6.1e-67;

Matches 194; Conservative 63; Mismatches 138; Indels 43; Gaps 13;

```

QY 30 TKK--VNTADQANRCTRNGKLPFTCAKAFVDFKARKQCLMPFNSSSGYKKEGHEFD 86
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 TKKQLGASIECAKCEEBE-EFTCRSFQYHSKEQOCVMAENRSSIYFRM--RDV 92
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 LYENKDYRNCLIGKGRSYKGTAVITKSGIKQPMSSMIPHEHSLPSSYRGKDLOENYC 146
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 LPEKMYLSECKNGKNGKRYGTAVSKTKGTLCQKMSAETPHKPRSPDENPEGLDQNYC 152
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 147 RNPGEEGPWCFTSNPEVRYEVCIDIPQCEVECTCKGSEYRGLMDHTESGKICQAWD 206
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 RNPNDGQPMWCYTTDPPEERFYCDIPEC-EDECMHSGENYDKISKTMGLGECQAWD 211
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 QTPRHAKFLPERYPDKGFDDNYCRNPDGQRPWCYTLDPHTRWEXCAIKTCADNTMNDT 266
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 QTPHAGTLPSEKFPKMLKKNYCRNPDGPRPWCFTTDPNKRWECDIPRCT-----TP 265
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 VPLE--TTECIQOGEGYGTAVTIWNGIPQQRWDSQYRHEHMTPEHFKCKDLRENYCR 324
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 PPSGPTQYCLKGTGENRQVAVATVSGHTCHGMSAQTPHTNRPENFPCKNLDENYCR 325
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 325 NPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GDCYRGNKN 367
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 326 NPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GDCYRGNKN 367
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 368 YMGNLSTGRSGLTCSMW-----DKMEDLHRIHFEWEPDASKLENYCRNPDDAAGP 419
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 YRGISSTTTGKCKCOSMSMTPHMHEKTPENF-----PNAAG-LTMYCANNPAD-KGP 435
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 420 WCYTGNPLIPWDCPISRCG 440
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 WCYTGNPLIPWDCPISRCG 440
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 441 TTDPSVRYEFCNLKCCSG 458
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

B30848
 plasmin (EC 3.4.21.7) precursor - rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
 C:Accession: B32869; B30848
 R:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
 J. Biol. Chem. 264, 5957-5965, 1989
 A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
 A:Reference number: A32869; MUID:89174660; PMID:2525643
 A:Accession: B32869
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-810 <TOM>
 A:Cross-references: UNIPROT:P12545; GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342272
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
 F:103-181/Domain: kringle homology <KR1>
 F:185-262/Domain: kringle homology <KR2>
 F:275-352/Domain: kringle homology <KR3>
 F:379-456/Domain: kringle homology <KR4>
 F:482-561/Domain: kringle homology <KR5>
 F:582-803/Domain: trypsin homology <TRY>

F:581-803/Domain: trypsin homology <TRY>
 F:49-73,53-61,103-181,124-164,152-176,185-262,168-316,206-245,234-257,275-352,296-335,32
 bonds: #status predicted
 F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 40.1%; Score 1044.5; DB 2; Length 810;
 Best Local Similarity 42.6%; Pred. No. 2.6e-66;

Matches 188; Conservative 74; Mismatches 128; Indels 51; Gaps 12;

```

QY 30 TKK--VNTADQANRCTRNGKLPFTCAKAFVDFKARKQCLMPFNSSSGYKKEGHEFD 86
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 TKKQLGASIECAKCEEBE-EFTCRSFQYHSKEQOCVMAENRSSIYFRM--RDV 92
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 LYENKDYRNCLIGKGRSYKGTAVITKSGIKQPMSSMIPHEHSLPSSYRGKDLOENYC 146
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 LPEKMYLSECKNGKNGKRYGTAVSKTKGTLCQKMSAETPHKPRSPDENPEGLDQNYC 152
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 147 RNPGEEGPWCFTSNPEVRYEVCIDIPQCEVECTCKGSEYRGLMDHTESGKICQAWD 206
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 RNPNDGQPMWCYTTDPPEERFYCDIPEC-EDECMHSGENYDKISKTMGLGECQAWD 211
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 QTPRHAKFLPERYPDKGFDDNYCRNPDGQRPWCYTLDPHTRWEXCAIKTCADNTMNDT 266
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 QTPHAGTLPSEKFPKMLKKNYCRNPDGPRPWCFTTDPNKRWECDIPRCT-----TP 265
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 VPLE--TTECIQOGEGYGTAVTIWNGIPQQRWDSQYRHEHMTPEHFKCKDLRENYCR 324
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 PPSGPTQYCLKGTGENRQVAVATVSGHTCHGMSAQTPHTNRPENFPCKNLDENYCR 325
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 325 NPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GDCYRGNKN 367
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 326 NPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GDCYRGNKN 367
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 368 YMGNLSTGRSGLTCSMW-----DKMEDLHRIHFEWEPDASKLENYCRNPDDAAGP 419
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 YRGISSTTTGKCKCOSMSMTPHMHEKTPENF-----PNAAG-LTMYCANNPAD-KGP 435
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 420 WCYTGNPLIPWDCPISRCG 440
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 WCYTGNPLIPWDCPISRCG 440
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 13

PLHM
 plasmin (EC 3.4.21.7) precursor [validated] - human
 N:Alternate names: plasminogen precursor [misnomer]
 N:Contains: angiotensin; microplasmin; plasminogen
 C:Species: Homo sapiens (man)
 C:Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
 C:Accession: A35229; A26646; A26738; A26735; A00929; A04627; A04625; A04
 R:Peterson, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
 J. Biol. Chem. 265, 6104-6111, 1990
 A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
 A:Reference number: A35229; MUID:90202879; PMID:2318848
 A:Accession: A35229
 A:Molecule type: DNA
 A:Residues: 1-810 <PER>
 A:Cross-references: UNIPROT:P00747; UNIPROT:Q9UB09; UNIPROT:Q9UM12; GB:J05286; GB:M34276
 A:Experimental source: leukocyte; lung fibroblast
 R:Margaret, N.; Bruno, L.; Portogallo, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; Ta
 Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
 A:Title: Definition of the transcription initiation site of human plasminogen gene in li
 A:Reference number: A35229; MUID:91097523; PMID:2266308
 A:Accession: A35229
 A:Status: translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-16 <MAL>
 A:Cross-references: GB:M62890; NID:G19092; PIDN:AAA36454.1; PID:G553613
 R:Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
 FEBS Lett. 213, 254-260, 1987
 A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
 A:Reference number: A26646; MUID:87162490; PMID:3030813
 A:Accession: A26646

A.Molecule type: mRNA
A.Residues: 1-471,'D',473-810 <FOR>
A.Cross-references: GB:X05199; NID:935530; PIDN:CAA28831.1; PID:935531
A.Experimental source: liver
R.Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A.Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A.Reference number: 145961; PMID:85023311; PMID:6148961
A.Accession: 162738
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 292-471,'D',473-810 <MAL2>
A.Cross-references: GB:X02922; NID:9190112; PIDN:AAA60124.1; PID:9387031
A.Accession: 184609
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 367-419 <MAL3>
A.Cross-references: GB:X02921; NID:9190110; PIDN:AAA60123.1; PID:9190111
R.Brundisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A.Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A.Reference number: S03735; PMID:81212097; PMID:7238497
A.Accession: S03735
A.Molecule type: protein
A.Residues: 20-71,'E',73-76 <BRU>
R.Sottrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A.Reference number: A00929
A.Accession: A00929
A.Molecule type: protein
A.Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOR>
R.Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A.Title: Primary structure of the B-chain of human plasmin.
A.Reference number: A04627; PMID:7725245; PMID:142009
A.Accession: A04627
A.Molecule type: protein
A.Residues: 581-810 <W11>
R.Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A.Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla
A.Reference number: A04625; PMID:75093329; PMID:122932
A.Accession: A04625
A.Molecule type: protein
A.Residues: 20-50,'Q',51-71,'E',73-85,87-100 <W12>
R.Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A.Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen tha
A.Reference number: A04626; PMID:76043692; PMID:126863
A.Accession: A04626
A.Molecule type: protein
A.Residues: 483-507,'E',509-604 <W13>
R.Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A.Title: The primary structure of human plasminogen. II. The histidine loop of human pla
A.Reference number: A93125; PMID:73149248; PMID:4694729
A.Contents: annotation; active site
R.Grodkopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A.Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A.Reference number: A92048; PMID:69234739; PMID:4420117
A.Contents: annotation; active site
R.Trexler, M.; Vail, Z.; Patchy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A.Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A.Reference number: A92382; PMID:82213905; PMID:6919539
A.Contents: annotation; omega-aminocarboxylic acid binding sites
R.Vail, Z.; Patchy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A.Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A.Reference number: A92459; PMID:85054794; PMID:6094526
A.Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site

R.Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehnle, S.; McCance, S.G.
J. Biol. Chem. 271, 29461-29467, 1996
A.Title: Kring domains of human angiotensin. Characterization of the anti-proliferati
A.Reference number: A58811; PMID:97067211; PMID:8910613
A.Contents: annotation
R.Lijnen, H.R.; Uguw, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A.Title: Generation of an angiotensin-like fragment from plasminogen by streptolysin-1
A.Reference number: A58812; PMID:9548733; PMID:9548733
A.Contents: annotation
R.Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A.Reference number: A51341; PDB:1PMK
A.Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R.Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A.Reference number: A51486; PDB:2PKA
A.Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R.Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A.Reference number: A51911; PDB:1PMK
A.Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R.Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A.Reference number: A52408; PDB:1PMK
A.Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R.Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A.Reference number: A65244; PDB:1CEA
A.Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R.Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A.Reference number: A65245; PDB:1CEB
A.Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R.Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A.Title: Crystal and molecular structure of human plasminogen kring 4 refined at 1.9
A.Reference number: A58819; PMID:92031502; PMID:1657148
A.Contents: annotation
R.Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A.Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmi
A.Reference number: A58818; PMID:92031503; PMID:1657149
A.Contents: annotation
R.de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M
Biochemistry 31, 270-279, 1992
A.Title: Crystal structure of the kring 2 domain of tissue plasminogen activator at 2
A.Reference number: A39483; PMID:92118803; PMID:1310033
A.Contents: annotation; X-ray crystallography, 2.4 angstroms
R.Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A.Reference number: A65980; PDB:1KRN
A.Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R.Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A.Reference number: A65803; PDB:1HPJ
A.Contents: annotation; conformation by (1)H-NMR, residues 103-181
R.Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A.Reference number: A65804; PDB:1HPK
A.Contents: annotation; conformation by (1)H-NMR, residues 103-181
R.Rejzante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A.Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasmin
A.Reference number: A58817; PMID:94237158; PMID:8181476
A.Contents: annotation; conformation by (1)H-NMR
C.Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU ;

Db 158 YCNPNDNDGWPWCYTTDDPKRYDCDIEPC-EDKCMHCSGENYEKIAWTMSGRCQAM 216
QY 205 DHQTPRHKFLPERYDCKGFDNDYCRNPGQPRPWCYTLDPHTRWYCAIKTGADTMD 264
Db 217 DSQSPAHAGTIPSPKPNKULKNMYCRNPQDEPRPWCYTLDPQKRWECDDPRLCT----- 270
QY 265 TDVPLE--TTECTIOGEGYRGTVNTIWNIGIPCORWDSQYPHEDMTPEHFCKDLRENY 322
Db 271 TTPSSGPRYQCLKGKGYGVAVTESGHTQORWSEQTPHKHRTPEHFCKDLRENY 330
QY 323 CRNPDSGSPWCYTLDPHTRWYCAIKTGADTMD-----GDCYRGNG 365
Db 331 CRNPDSGSPWCYTLDPHTRWYCAIKTGADTMD-----GDCYRGNG 389
QY 366 KNYMGLSOTRSGTSCMDKXNEDLHRIHFWBPDASKLNENYCRNPDDAHGPMWCY 422
Db 390 QSTRGTSSTTITRKCQSSWMP--HRLKTPENYPNAG-LTMNYCRNPDDAHGPMWCY 445
QY 423 TGNPLIPWDYCPISRC 438
Db 446 TTDPRVMEFCNLKCC 461

RESULT 15

T18518
apolipoprotein(a) - western European hedgehog (fragment)
C:Species: Erinaceus europaeus (western European hedgehog)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18518
J. Lam, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein
A:Reference number: I46259; PMID:96025778; PMID:7592597
A:Accession: T18518
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2869 <LAW>
A:Cross-references: UNIPROT:Q28398; EMBL:U33170; NID:g1046358; PID:g1046359; PIDN:AA485
A:Experimental source: liver
C:Comment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, contains apolipoprotein(a).

Query Match 36.2%; Score 943.5; DB 2; Length 2869;
Best Local Similarity 45.6%; Pred. No. 1.4e-58;
Matches 173; Conservative 51; Mismatches 126; Indels 29; Gaps 10;

QY 89 ENKDYTRNCIIGKRSYKGTSTTKSGIKCQWSSMTPEHSPFLPSSYRGKDLQENYCRN 148
Db 2490 EPLDLVDQCLEGTGENTRGMAVTASGHTCQWRREQSPHSHSRTPENYPTKNLVGNVCRN 2549
QY 149 PRGEGGPMCFNSNPRVYEVCDIPQCS-----VEGMCNGESYRGLMDHTS 197
Db 2550 PGEV-APWCYTTNSAMRWECISIPACESPTPEHLVPEQCLENGENYQGNMAITVS 2608
QY 198 GKICQKMDQTPHRAHFLPERYDCKGFDNDYCRNPGQPRPWCYTLDPHTRWYCAIKTC 257
Db 2609 GQPCQGRKQTPHRAHFLPERYDCKGFDNDYCRNPGQPRPWCYTLDPHTRWYCAIKTC 2668
QY 258 ADNTMNDTVPLETTCTIOGEGYRGTVNTIWNIGIPCORWDSQYPHEDMTPEHFCKDX 317
Db 2669 -ESSSPTEPMIIPDQCLRGTEGENTRGMAVTASGHTCQWRREQSPHSHSRTPENYPTKN 2727
QY 318 LRENYCRNPDSGSPWCYTLDPHTRWYCAIKTGADTMD-----GDCYRGNG 369
Db 2728 LRENYCRNPDSGSPWCYTLDPHTRWYCAIKTGADTMD-----GDCYRGNG 389
QY 370 GNLSTQTRSGTSCMDKXNEDLHRIHFWBPDASKLNENYCRNPDDAHGPMWCYGNP 426
Db 2787 GNLSTQTRSGTSCMDKXNEDLHRIHFWBPDASKLNENYCRNPDDAHGPMWCYGNP 446
QY 427 LIPWDYCPISRCGDTTPT 445

Db 2843 AVRWEYCSIPTCESPTTPT 2861
Search completed: October 3, 2005, 05:59:59
Job time: 41.225 secs

THIS PAGE BLANK (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:29 ; Search time 65.3656 Seconds
(without alignments)
3501.834 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604
Sequence: 1 EKKRRNTLHEFKSAKXTLLI.....IPWDYCPISRCGGTPTTIV 447

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_03:*
2: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2601	99.9	728	1 HGF_HUMAN	P14210 homo sapien
2	2462	94.5	730	2 Q867B7	Q867B7 canis famil
3	2456	94.3	730	2 Q76BS1	Q76BS1 bos taurus
4	2452	94.2	728	2 Q9BH09	Q9BH09 felis silve
5	2409	92.5	728	2 Q8C9C5	Q8C9C5 mus musculu
6	2408	92.5	728	1 HGF_MOUSE	Q08048 mus musculu
7	2401	92.2	728	1 HGF_RAT	P17945 rattus norv
8	2052	78.8	726	2 Q90978	Q90978 gallus gall
9	1864.5	71.6	720	2 Q91402	Q91402 xenopus. he
10	1705.5	65.5	449	2 Q788Q2	Q788Q2 gallus gall
11	1214	46.6	449	2 Q6GR14	Q6GR14 xenopus lae
12	1209.5	46.4	704	2 Q90865	Q90865 gallus gall
13	1208	45.4	716	2 Q91691	Q91691 xenopus lae
14	1180.5	45.3	711	1 HGF_HUMAN	P26927 homo sapien
15	1176	45.2	717	2 P70006	P70006 xenopus lae
16	1172	44.9	709	2 Q90ZM6	Q90ZM6 brachydant
17	1168.5	44.9	709	2 Q7ZTN9	Q7ZTN9 xenopus lae
18	1163.5	44.7	567	2 Q13Z08	Q13Z08 homo sapien
19	1144	43.9	716	2 P70521	P70521 rattus norv
20	1125	43.2	716	1 HGF_MOUSE	P26928 mus musculu
21	1125	43.2	716	2 Q91XG8	Q91XG8 mus musculu
22	1089	41.8	818	2 Q6PBA6	Q6PBA6 brachydant
23	1085.5	41.7	812	1 PLMN_RAT	O01177 rattus norv
24	1080.5	41.5	806	1 PLMN_MOUSE	O18783 macropus eu
25	1066	40.9	790	1 PLMN_PIG	P06867 sus scrofa
26	1064	40.9	466	2 Q6TC10	Q6TC10 sus musculu
27	1064	40.9	812	1 PLMN_MOUSE	P20918 mus musculu
28	1053.5	40.5	810	1 PLMN_ERIEU	Q29485 erinaceus e
29	1044.5	40.1	810	1 PLMN_MACMU	P12545 macaca mula
30	1042	40.0	810	1 PLMN_HUMAN	P00747 homo sapien
31	1040	39.9	215	2 O42341	O42341 gallus gall

32	1038	39.9	812	1 PLMN_BOVIN	P06668 bos taurus
33	1032.5	39.7	759	2 Q7TP84	Q7TP84 rattus norv
34	985	37.8	359	2 Q8MMR1	Q8MMR1 canis famil
35	966.5	37.1	648	2 Q9H1V4	Q9H1V4 homo sapien
36	943.5	36.2	2869	2 Q28398	Q28398 erinaceus e
37	896	34.4	211	2 Q55027	Q55027 mus musculu
38	879	33.8	4548	1 APOA_HUMAN	P08519 homo sapien
39	843.5	32.4	1420	1 APOA_MOUSE	P14417 macaca mula
40	617	23.7	162	2 Q8C4E2	Q8C4E2 mus musculu
41	615	23.6	109	2 Q9N1B8	Q9N1B8 ovis aries
42	444.5	17.1	454	2 Q46506	Q46506 papio hamad
43	439	16.9	75	2 Q9BGN9	Q9BGN9 bos taurus
44	408.5	15.7	325	1 PLMN_PETUMA	P33574 petromyzon
45	406	15.6	429	2 Q8AVB0	Q8AVB0 brachydant

ALIGNMENTS

RESULT 1	ID	HGF_HUMAN	STANDARD:	PRT:	728 AA.
AC	P14210	002935	Q13494	Q14519	Q8RC62
DT	01-JUN-1990	(Rel. 13, Created)			
DT	01-AUG-1991	(Rel. 19, Last sequence update)			
DT	25-OCT-2004	(Rel. 45, Last annotation update)			
DE	Hepatocyte growth factor precursor (Scatter factor) (SF)				
DE	(Hepatopoietin A).				
GN	Name=HGF; Synonyms=HPTA;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=91340155; PubMed=1831432; DOI=10.1016/0378-1119(91)90080-U;				
RA	Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;				
RT	"Organization of the human hepatocyte growth factor-encoding gene.";				
RL	Gene 102:213-219(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Placenta;				
RX	MEDLINE=89392017; PubMed=2528952;				
RA	Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,				
RA	Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,				
RT	Gohda E., Daikuhara Y., Kitamura N.;				
RT	"Molecular cloning and sequence analysis of cDNA for human hepatocyte				
RL	growth factor.";				
RL	Biochem. Biophys. Res. Commun. 163:967-973(1989).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Leukocyte;				
RX	MEDLINE=91025062; PubMed=2145836;				
RA	Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T., Asami O.,				
RA	Hagiya M., Nakamura T., Shimizu S.;				
RT	"Isolation and expression of cDNA for different forms of hepatocyte				
RL	growth factor from human leukocyte.";				
RL	Biochem. Biophys. Res. Commun. 172:321-327(1990).				
RN	[4]				
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 55-73 AND 495-520.				
RC	TISSUE=Liver;				
RX	MEDLINE=9006676; PubMed=2511289; DOI=10.1038/342440a0;				
RA	Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,				
RA	Sugimura A., Tashiro K., Shimizu S.;				
RT	"Molecular cloning and expression of human hepatocyte growth factor.";				
RL	Nature 342:440-443(1989).				
RN	[5]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	TISSUE=Placenta;				
RX	MEDLINE=91200041; PubMed=1826653;				
RA	Miyazawa K., Kitamura A., Naka D., Kitamura N.;				
RT	"An alternatively processed mRNA generated from human hepatocyte				
RL	growth factor gene.";				

- RL Eur. J. Biochem. 197;15-22(1991).
- RN [6]
- RA SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 583-592.
- RP TISSUE=lung fibroblast;
- RC PubMed=1824873;
- RA Rubin J.S., Chan A.M.-L., Bottaro D.P., Burgess W.H., Taylor W.G.,
- RA Cecch A.C., Hirschfield D.W., Wong J., Miki T., Finch P.W.,
- RA Aaronson S.A.;
- RT "A broad-spectrum human lung fibroblast-derived mitogen is a variant
- RT of hepatocyte growth factor";
- RL Proc. Natl. Acad. Sci. U.S.A. 88:415-419(1991).
- RN [17]
- RA SEQUENCE FROM N.A. (ISOFORM 1).
- RP TISSUE=embryonic fibroblast;
- RC MEDLINE=91334393; PubMed=1831266;
- RA Weidner K.M., Arkaki N., Hartman G., Vandekerckhove J., Weingart S.,
- RA Riederer H., Fomatech C., Tsubouchi H., Hishida T., Dalkuhara Y.,
- RA Birnmeier W.;
- RT "Evidence for the identity of human scatter factor and human
- RT hepatocyte growth factor";
- RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
- RN [18]
- RA SEQUENCE FROM N.A. (ISOFORM 4).
- RP PubMed=1720571;
- RA Chan A.M.-L., Rubin J.S., Bottaro D.P., Hirschfield D.W., Chened M.,
- RA Aaronson S.A.;
- RT "Identification of a competitive HGF antagonist encoded by an
- RT alternative transcript";
- RL Science 254:1382-1385(1991).
- RN [19]
- RA SEQUENCE FROM N.A. (ISOFORM 2), AND MUTAGENESIS OF ARG-494.
- RP MEDLINE=93087571; PubMed=1280830;
- RA Hartman G., Nalidini L., Weidner K.M., Sachs M., Vigna E.,
- RA Comoglio P.M., Birnmeier W.;
- RT "A functional domain in the heavy chain of scatter factor/hepatocyte
- RT growth factor binds the c-Met receptor and induces cell dissociation
- RT but not mitogenesis";
- RL Proc. Natl. Acad. Sci. U.S.A. 89:11574-11578(1992).
- RN [110]
- RA SEQUENCE FROM N.A. (ISOFORM 6).
- RP MEDLINE=96278713; PubMed=8662798; DOI=10.1074/jbc.271.22.13110;
- RA Cioce V., Csaky K.G., Chan A.M.-L., Bottaro D.P., Taylor W.G.,
- RA Jensen R., Aaronson S.A., Rubin J.S.;
- RT "Hepatocyte growth factor (HGF)/NK1 is a naturally occurring
- RT HGF/scatter factor variant with partial agonist/antagonist activity";
- RL J. Biol. Chem. 271:13110-13115(1996).
- RN [111]
- RA SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS LYS-304 AND TYR-330.
- RP Riederer H., Livingston R.J., Daniels M.R., Chung M.-W.,
- RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
- RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
- RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department
- RT of Genome Sciences, Seattle, WA (URL: <http://egp.gs.washington.edu>).";
- RL Submitted (FEB-2003) to the EMBL/GenBank/DBD databases.
- RN [112]
- RA SEQUENCE FROM N.A.
- RP MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
- RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
- RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
- RA Wylie K., Sekhon M., Becker M.C., O'Laughlin W.D., Schaller M.E.,
- RA Fewell G.A., Delenanthu K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
- RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
- RA VanHunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
- RA Ozeraky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
- RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
- RA Kozlowski-Reilly A., Leonard S., Kohlfing T., Rock S.M.,
- RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Stromatt C.,
- RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
- RA Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Splich L.,
- RA Bieri T., Nelson J.O., Berkowicz N., Wohlmann P.E., Cook L.L.,
- RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
- RA Clifton S.W., Chisoe S.D., Marra M.A., Raymond C., Haugen E.,
- RA Gillett W., Zhou Y., James R., Phelps K., Iadonato S., Hubb K.,
- RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
- RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
- RA Bailey J.A., Portnoy M.E., Torrents D., Chirwalla A.T., Gish W.R.,
- RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
- RA Waterston R.H., Wilson R.K.;
- RT "The DNA sequence of human chromosome 7";
- RL Nature 424:157-164(2003).
- RN [113]
- RA SEQUENCE FROM N.A. (ISOFORMS 5 AND 6).
- RP MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
- RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
- RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
- RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
- RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
- RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
- RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
- RA Brownstein M.J., Uesli T.B., Toshilyki S., Carantini P., Prange C.,
- RA Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
- RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
- RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
- RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
- RA Fahy J., Hellon E., Kettleman M., Madan A., Rodriguez S., Sanchez A.,
- RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
- RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
- RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
- RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
- RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
- RT "Generation and initial analysis of more than 15,000 full-length human
- RT and mouse cDNA sequences";
- RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
- RN [114]
- RA SEQUENCE OF 1-208 AND 249-695 FROM N.A. (ISOFORM 1).
- RP MEDLINE=91329192; PubMed=1832556;
- RA Miyazawa K., Kitamura A., Kitamura N.;
- RT "Structural organization and the transcription initiation site of the
- RT human hepatocyte growth factor gene";
- RL Biochemistry 30:9170-9176(1991).
- RN [115]
- RA SIGNAL SEQUENCE CLEAVAGE SITE.
- RP MEDLINE=91207365; PubMed=1826837;
- RA Yoshiyama Y., Arkaki N., Naka D., Takahashi K., Hiroo S., Kondo J.,
- RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
- RA Hishida T., Dalkuhara Y.;
- RT "Identification of the N-terminal residue of the heavy chain of both
- RT native and recombinant human hepatocyte growth factor";
- RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
- RN [116]
- RA CARBOHYDRATE-LINKAGE SITE THR-476.
- RP MEDLINE=93129192; PubMed=182348;
- RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
- RA Nakamura T., Shimizu S.;
- RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
- RT on the alpha chain";
- RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
- RN [117]
- RA MUTAGENESIS.
- RP MEDLINE=92331602; PubMed=1321034;
- RA Loker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
- RA Baker J.B., Godowski P.J.;
- RT "Structure-function analysis of hepatocyte growth factor:
- RT identification of variants that lack mitogenic activity yet retain
- RT high affinity receptor binding";
- RL EMBO J. 11:2503-2510(1992).
- RN [118]
- RA STRUCTURE BY NMR OF 31-127.
- RP MEDLINE=9815423; PubMed=9493272; DOI=10.1016/S0969-2126(98)00012-4;
- RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
- RA Rubin J.S., Bottaro D.P., Byrd R.A.;
- RT "The solution structure of the N-terminal domain of hepatocyte growth
- RT factor reveals a potential heparin-binding site";
- RL Structure 6:109-116(1998).
- RN [119]
- RA X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.

RX MEDLINE=99036858; PubMed=9817840; DOI=10.1016/S0969-2126(98)00138-5;
RA Ultsch M., Lokker N.A., Godowski P.J., de Vos A.M.;

Query Match 99.9%; Score 2601; DB 1; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.1e-186;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ERKRRNTIHEFKSAKTLTIKIDPAIKITKKNVTADQCANRCKRNGKLPFTCKAFVFDK 60
DB 32 QRRRRNTIHEFKSAKTLTIKIDPAIKITKKNVTADQCANRCKRNGKLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDIYIRNCIIIGKGSYKGTVSITSGIKCOP 120
DB 92 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDIYIRNCIIIGKGSYKGTVSITSGIKCOP 151
QY 121 WSMIPIHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFNSPEVYEVCDIPQCESEVFC 180
DB 152 WSMIPIHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFNSPEVYEVCDIPQCESEVFC 211
QY 181 MTCNGESYRGKDLMDHTESGKICQRMHQTPRHKFLPERYPDKGFDNNYCRNPDGQRPWC 240
DB 212 MTCNGESYRGKDLMDHTESGKICQRMHQTPRHKFLPERYPDKGFDNNYCRNPDGQRPWC 271
QY 241 YTLDPTRWEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGTVNTIWNIGIPCORWDS 300
DB 272 YTLDPTRWEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGTVNTIWNIGIPCORWDS 331
QY 301 OYPHEHDMTPENPKCKDLKENYCRNPDGSESPMCFITTDNIRVYCSQIPNCMESHGQDC 360
DB 332 QYPHEHDMTPENPKCKDLKENYCRNPDGSESPMCFITTDNIRVYCSQIPNCMESHGQDC 391
QY 361 YGNGKNYMGNLISQTRSGJLTCGSMWCKNMDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420
DB 392 YGNGKNYMGNLISQTRSGJLTCGSMWCKNMDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCGDTTPTIV 478

RESULT 2
ID Q867B7 PRELIMINARY; PRT; 730 AA.
AC Q867B7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hepatocyte growth factor.
GN Name=hgf;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle;
RX MEDLINE=22844761; PubMed=12963274; DOI=10.1016/S0165-2427(03)00118-1;
RA Miyake M., Saze K., Yaguchi T., Wang J., Suzuta Y., Haga Y.,
RA Takahashi S.Y., Yamamoto Y., Iwabuchi S.;
RT "Canine hepatocyte growth factor: molecular cloning and
RT characterization of the recombinant protein."
RL Vet. Immunol. Immunopathol. 95:135-143(2003).
RN 12)
RP SEQUENCE FROM N.A.
RA Liao A.T., Chien M.B., London C.A.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family 51.
DR EMBL; AB090353; BAC57560.1; -.
DR HSSP; P14210; IGMN.
DR MEROPS; S01.982; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
```

```
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR011359; HGF_MST1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00051; Kringle_4.
DR Pfam; PF00024; PAN_1.
DR Pfam; PF00089; Trypsin_1.
DR PIRSF; PIRSF001152; HGF_MST1_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00473; PAN_AP_1.
DR SMART; SM00020; TRYP_SPE_1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00948; PAN_1.
DR PROSITE; PS0240; TRYPSIN_DOM_1.
KW Hydrolyase; Kringle; Protease; Serine protease.
SQ SEQUENCE 730 AA; 83398 MW; 4E83F9EAB8FF6DB1 CRC64;

Query Match 94.5%; Score 2462; DB 2; Length 730;
Best Local Similarity 93.1%; Pred. No. 2.9e-176;
Matches 416; Conservative 19; Mismatches 12; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLTIKIDPAIKITKKNVTADQCANRCKRNGKLPFTCKAFVFDK 60
DB 32 QRRRRNTIHEFKSAKTLTIKIDPAIKITKKNVTADQCANRCKRNGKLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDIYIRNCIIIGKGSYKGTVSITSGIKCOP 120
DB 92 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDIYIRNCIIIGKGSYKGTVSITSGIKCOP 151
QY 92 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDIYIRNCIIIGKGSYKGTVSITSGIKCOP 151
DB 92 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDIYIRNCIIIGKGSYKGTVSITSGIKCOP 151
QY 121 WSMIPIHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFNSPEVYEVCDIPQCESEVFC 180
DB 152 WSMIPIHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFNSPEVYEVCDIPQCESEVFC 211
QY 181 MTCNGESYRGKDLMDHTESGKICQRMHQTPRHKFLPERYPDKGFDNNYCRNPDGQRPWC 240
DB 212 MTCNGESYRGKDLMDHTESGKICQRMHQTPRHKFLPERYPDKGFDNNYCRNPDGQRPWC 271
QY 241 YTLDPTRWEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGTVNTIWNIGIPCORWDS 300
DB 272 YTLDPTRWEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGTVNTIWNIGIPCORWDS 331
QY 301 OYPHEHDMTPENPKCKDLKENYCRNPDGSESPMCFITTDNIRVYCSQIPNCMESHGQDC 360
DB 332 QYPHEHDMTPENPKCKDLKENYCRNPDGSESPMCFITTDNIRVYCSQIPNCMESHGQDC 391
QY 361 YGNGKNYMGNLISQTRSGJLTCGSMWCKNMDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420
DB 392 YGNGKNYMGNLISQTRSGJLTCGSMWCKNMDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCGDTTPTIV 478

RESULT 3
ID Q76BS1 PRELIMINARY; PRT; 730 AA.
AC Q76BS1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hepatocyte growth factor.
GN Name=hgf;
OS Bos taurus (Bovine).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxId=9913;
RN (1)
RP SEQUENCE FROM N.A.
RA Yamaji D., Kimura K., Watanabe A., Makondo K., Salto M.,
RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family 11.
DR EMBL; AB110822; BAD02475.1; -.
DR HSBP; P08519; IKIV.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR011359; HGF_MST1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PIRSF; PIRSF001152; HGF_MST1; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR000395; Kringle; 4.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS50948; PAN; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR Hydrolase; Kringle; Protease; Serine protease.
KW Hydrolyase; Kringle; Protease; Serine protease.
SQ SEQUENCE 730 AA; 83357 MW; E39D653B0A85F49B CRC64;

Query Match 94.3%; Score 2456; DB 2; Length 730;
Best Local Similarity 92.6%; Pred. No. 8.2e-176;
Matches 414; Conservative 22; Mismatches 11; Indels 0; Gaps 0;

QY 1 EKKRNTIHEFKSAKTLIKIDPALKIKTKKVNATADQCARCTRNKGLPTCKAFVFDK 60
DB 32 QKKRNTLHEFKSAKTLIKEDPLKIKTKKNTADQCARCIRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKCOP 120
DB 92 ARKRCIMFPNMSGVKKFGEHFDLYENKDYIRNCIIIGKGSYKGVSTIKSGIKCOP 151
QY 121 WSSMTPIHESFLPSSYRGKDLQENYCRNRGEGGWCMTSNPEVYEVCDIPQSEVVC 180
DB 152 WSSMTPIHESFLPSSYRGKDLQENYCRNRGEGGWCMTSNPEVYEVCDIPQSEVVC 211
QY 181 MTCNGESYRGKDLMDHTESGKICORMDHOTPHRKFLPERYPDGFDDNYCRNPDGQPRPVC 240
DB 212 MTCNGESYRGKDLMDHTETGKICQRMDHOTPHRKFLPERYPDGFDDNYCRNPDGQPRPVC 271
QY 241 YLUDPHTRWEYCAIKTCADNTMNDTDLVLETTETCIQGGEGYRGVNTIWIANGIPCOQRWDS 300
DB 272 YLUDPHTRWEYCAIKTCADNTMNDTDLVLETTETCIQGGEGYRGVNTIWIANGIPCOQRWDS 331
QY 301 QYRPHEDMTPEPKCKDLRENYCRNPDGSESPWCFTTDINIVGYSQIPNCDMSGQDC 360
DB 332 QYRPHEDMTPEPKCKDLRENYCRNPDGSESPWCFTTDINIVGYSQIPNCDMSGQDC 391
QY 361 YGNGKNYVGNLSQTSRGLTCSWMDKGMEDLHRHIFWEPDASKLNENYCRNPDGDAHGMW 420
DB 392 YGNGKNYVGNLSQTSRGLTCSWMDKGMEDLHRHIFWEPDASKLNENYCRNPDGDAHGMW 451
QY 421 CYTGNPLIPMDYCPISRCGDTPTTIV 447

DB 452 CYTGNPLIPMDYCPISRCGDTPTTIV 478

RESULT 4
ID Q9BH09 PRELIMINARY; PRT; 728 AA.
AC Q9BH09;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hepatocyte growth factor HGF.
GN Name=HGF;
OS Fetus silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxId=9685;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kobayashi Y., Nakamura N., Ishizaka T., Masuda K., Ohno K.,
RA Tsujimoto H.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Miyake M., Yaguchi T., Saze K., Suzuta Y., Wang J., Okazaki M.,
RA Haga Y., Yamamoto Y., Takahashi S., Iwabuchi S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family 11.
DR EMBL; AB046610; BAB21499.1; -.
DR EMBL; AB080187; BAC10545.1; -.
DR HSBP; P14210; IGNN.
DR MEROPS; S01.982; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR011359; HGF_MST1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PIRSF; PIRSF001152; HGF_MST1; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR000395; Kringle; 4.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS50948; PAN; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR Hydrolase; Kringle; Protease; Serine protease.
KW Hydrolyase; Kringle; Protease; Serine protease.
SQ SEQUENCE 728 AA; 83067 MW; 8D7F4A333D1E190A CRC64;

Query Match 94.2%; Score 2452; DB 2; Length 728;
Best Local Similarity 92.6%; Pred. No. 1.6e-175;
Matches 414; Conservative 21; Mismatches 12; Indels 0; Gaps 0;

QY 1 EKKRNTIHEFKSAKTLIKIDPALKIKTKKVNATADQCARCTRNKGLPTCKAFVFDK 60
DB 30 QKKRNTLHEFKSAKTLIKEDPLKIKTKKNTADQCARCIRNKGLPTCKAFVFDK 89
QY 61 ARKQCLMPFPNMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKCOP 120
DB 90 ARKRCIMFPNMSGVKKFGEHFDLYENKDYIRNCIIIGKGSYKGVSTIKSGIKCOP 149
QY 121 WSSMTPIHESFLPSSYRGKDLQENYCRNRGEGGWCMTSNPEVYEVCDIPQSEVVC 180

```
Db 150 MNSMIPHEHSFLPSSYRGKDLQENYCNPRGEGGPMCFSTNSEVRYEVCDDIQCSBEVC
Qy 181 MTCNGESYRGLMDHTSGKICQRMWDQTPRRHKFLPERYDKGFDDNYCGRPDPGQRPWC
Db 210 MTCNGESYRGLMDHTSGKICQRMWDQTPRRHKFLPERYDKGFDDNYCGRPDPGQRPWC
Qy 241 YTLDPTRMEYCAIKTCADNTMNDTDPVLETTETCICQGGEGYRGVTIWNIGPICQRMWS
Db 270 YTLDPTRMEYCAIKTCADNTMNDTDPVLETTETCICQGGEGYRGVTIWNIGPICQRMWS
Qy 301 QYFHEHDMTPENKCKCDLRNRYCRNPDGSGSPWCFTTDPNIRYGCQIIPNCMSHGQDC
Db 330 QYFHOHDIPTENPKCKDRLRNFRCNPDGASPMCFPTTDPNIRYGCQIIPKCVSSGQDC
Qy 361 YRGNKGYKMGNTLSQTSRGLTCGSMWDMKMDLHHIFWEPASKLINENYCRNPDGAGPW
Db 390 YRGNKGYKMGNTLSQTSRGLTCGSMWDMKMDLHHIFWEPASKLINENYCRNPDGAGPW
Qy 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 450 CYTGNPLIPWDYCPISRCBGDTTPTIV 476

RESULT 5
ID Q8C9G5 PRELIMINARY; PRT; 728 AA.
AC Q8C9G5;
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
DE library, clone:A630060C16 product:hepatocyte growth factor, full
DE insert sequence.
GN Name:Hgf;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
```

```
RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Katsunai T., Taahiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohata E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family S1.
DR EMBL: AK042121; BAC31175.1; -.
DR HSSP: P14210; IGMN.
DR WGD; MG1:96079; Hgf.
DR GO; GO:0008283; P:cell proliferation; IDA.
DR GO; GO:0008282; P:cellular morphogenesis; IDA.
DR InterPro: IPR011359; HGF_MST1.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; Trypsin; 1.
DR PIRSF: PIRSF001152; HGF_MST1; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRY_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS00070; KRINGLE_2; 4.
DR PROSITE: PS00948; PAN; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
KW Hydrolase; Kringle; protease; Serine protease.
SQ SEQUENCE 728 AA; 82990 MW; 8D5258DF3BCF3545 CRC64;

Query Match 92.5%; Score 2409; DB 2; Length 728;
Best local Similarity 91.1%; Pred. No. 2,7e-172;
Matches 407; Conservative 21; Mismatches 19; Indels 0; Gaps 0;

Qy 1 EKKRRNTIHEPKSAKATLTIKIDPALKIKTKKVTADQCANRCTRNGSLPFTCKAFVDPK
Db 33 OKRRNTIHEPKSAKATLTIKIDPALKIKTKKVTADQCANRCTRNGSLPFTCKAFVDPK
Qy 61 ARKQCLMPFNSMSSGVKKEFGHEFDLVENKYYINCIIGKRSYKGYTISITKSGIKCOP
Db 93 SRKRCWYFPNMSGKVGKGFGEHFDLVENKYYINCIIGKGSYKGYTISITKSGIKCOP
Qy 121 MNSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNSEVRYEVCDDIQCSBEVC
Db 153 MNSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNSEVRYEVCDDIQCSBEVC
Qy 181 MTCNGESYRGLMDHTSGKICQRMWDQTPRRHKFLPERYDKGFDDNYCGRPDPGQRPWC
```

```
Db 213 MTCNSESVRGPMHDTESGKTCQRMDOQTPEHRHAKFLPERKPDGKFPDNNYCRNPDGKPRPMC 272
Oy 241 YTLDDHTWMEYCAITTCADNTMNDVDVPLETTECTIOGEGRGVNTIMNGICPORRDS 300
Db 273 YTLDDHTWMEYCAITTCASAVENETDVPETTECTIOGEGRGVNTIMNGICPORRDS 332
Oy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGCSOI PNCDMSHGQDC 360
Db 333 QYPHGHDTPEPFKCKDLRENYCRNPDGSESPWCTTDPNIRVGCSOI PNCDVSSGQDC 392
Oy 361 YRGNKNYMGNISQTRSGITCSMDKNMEDLHRHIFWEEDASKLNENYCRNPDGDAHGPW 420
Db 393 YRGNKNYMGNISKTRSGITCSMDKNMEDLHRHIFWEEDASKLNENYCRNPDGDAHGPW 452
Oy 421 CYTGNPLIPMDYCPISRCGGDTTPITV 447
Db 453 CYTGNPLIPMDYCPISRCGGDTTPITV 479
```

RESULT 6

HGF_MOUSE STANDARD; PRT; 728 AA.

```
ID_HGF_MOUSE 008048; Q61662; Q64007;
AC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
   (Hepatopoietin A).
GN Name=Hgf;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), AND SEQUENCE OF 496-504.
RC TISSUE=Mammary fibroblast;
RX MEDLINE=94183257; PubMed=815822;
RA Sasaki M., Nishio M., Sasaki T., Enami J.;
RT "Identification of mouse mammary fibroblast-derived mammary growth
   factor as hepatocyte growth factor.";
RL Biochem. Biophys. Res. Commun. 199; 772-779 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94363381; PubMed=8081873;
RA Lee C.C., Kozak C.A., Yamada K.M.;
RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter
   factor gene.";
RL Cell Adhes. Commun. 1; 101-111 (1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94060105; PubMed=8241272; DOI=10.1016/0167-4781(93)90159-B;
RA Liu Y., Michalopoulos G.K., Zarnegar R.;
RT "Molecular cloning and characterization of cDNA encoding mouse
   hepatocyte growth factor.";
RL Biochim. Biophys. Acta 1216; 299-303 (1993).
CC -I- FUNCTION: HGF is a potent mitogen for mature parenchymal
   hepatocyte cells, seems to be an hepatotrophic factor, and acts as
   growth factor for a broad spectrum of tissues and cell types. It
   has no detectable protease activity.
CC -I- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
   disulfide bond.
CC -I- ALTERNATIVE PRODUCTS:
   Event=Alternative splicing; Named isoforms=2;
   Name=Long;
   IsoId=Q08048-1; Sequence=Displayed;
   Name=Short;
   IsoId=Q08048-2; Sequence=VSP_005408;
CC -I- SIMILARITY: Belongs to the peptidase S1 family. Plasminogen
   subfamily.
CC -I- SIMILARITY: Contains 4 kringle domains.
```

-I- SIMILARITY: Contains 1 PAN domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

```
DR EMBL; D10212; BAA01064.1; -
DR EMBL; D10213; BAA01065.1; -
DR EMBL; S71816; AAB31855.1; -
DR EMBL; X72307; CAA51054.1; ALT_INIT.
DR PIR; JC2117; A60185.
DR HSSP; P14210; 1BHT.
DR MEROPS; S01.982; -.
DR MGD; MGI:96079; Hgf.
DR GO; GO:0008283; P:cell proliferation; IDA.
DR GO; GO:000902; P:cellular morphogenesis; IDA.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PIRSF; PIRSF001152; HGF_MST1; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SMO0130; KR; 4.
DR SMART; SMO0473; PAN_AP; 1.
DR SMART; SMO0020; TRY_P_Ser; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00948; PAN; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KM Alternative splicing; Direct protein sequencing; Glycoprotein;
KM Growth factor; Kringle; Pyroglutamate carboxylic acid; Repeat;
KW Serine protease homolog; Signal.
FT SIGNAL 1 32
FT CHAIN 33 495
FT CHAIN 496 728
FT MOD_RES 33 33
FT DOMAIN 38 124
FT DOMAIN 129 207
FT DOMAIN 212 289
FT DOMAIN 306 384
FT DOMAIN 392 470
FT DOMAIN 496 728
FT DISULFID 71 97
FT DISULFID 75 85
FT DISULFID 488 607
FT CARBOHYD 295 295
FT CARBOHYD 403 403
FT CARBOHYD 569 569
FT CARBOHYD 556 556
FT VARSPPLIC 163 167
FT CONFLICT 344 344
FT CONFLICT 479 479
FT CONFLICT 564 564
SQ SEQUENCE 728 AA; 82944 MW; A0381FC497534328 CRC64;
```

Query Match 92.5%; Score 2408; DB 1; Length 728;

Best Local Similarity 91.1%; Pred. No. 3,2e-172;

Oy 1 ERKRNTIHEPKSAKTLIKIDPALKIKIKYKVVNTADQCANCRTRNKGLPFTCAAFVFDK 60
Matches 407; Conservative 21; Mismatches 19; Indels 0; Gaps 0;

```

Db 33 OKRRNTLHEFKKSAKTLTKEDPLKIKTKKVNNSADECANRCIRNKGFTTCKAFVFDK 92
Qy 61 ARKQCLMPFPNMSGCVKKEFGHEPDLVKNKYIRNCIIIGKGRSYGTVSITKSGIKCOP 120
Db 93 SKRRCYWPFPNMSGCVKKGFGHEPDLVKNKYIRNCIIIGKGSYKGTVSITKSGIKCOP 152
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCESEVC 180
Db 153 WMSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCESEVC 212
Qy 181 MTCNGESTYGLMDHTESGKICORPDHQTFRHKFLPERYPDKGPDNNYCRNPDGQRPWC 240
Db 213 MTCNGESTYRGPMHTESGKICORPDQTPHRHKFLPERYPDKGPDNNYCRNPDGKRPWC 272
Qy 241 YTLDDHTREVCATKTCANTMNDVPLETTECTOCGEGYRGATNTIWNIGPCQRMWS 300
Db 273 YTLDDHTREVCATKTCANSAVETDVPWETTECTOCGEGYRGATNTIWNIGPCQRMWS 332
Qy 301 QYRPHDMTPENFKCKDLRENYCRNPDGSESPWCTTDDNIRVYCSQIPNCDMSHGDC 360
Db 333 QYRPHMDITPENFKCKDLRENYCRNPDGSESPWCTTDDNIRVYCSQIPKCDVSSGDC 392
Qy 361 YRNGKNYNGNLSQTRSGITCSMMDKXMDLHRHIFWEPDASKLNENYCRNPDGDAHGW 420
Db 393 YRNGKNYNGNLSKTRSGITCSMMDKXMDLHRHIFWEPDASKLNENYCRNPDGDAHGW 452
Qy 421 CYTGNPLIPMDYCPISRCGSDTPTTV 447
Db 453 CYTGNPLIPMDYCPISRCGSDTPTTV 479

RESULT 7
HGF_RAT STANDARD; PRT; 728 AA.
ID PI7945;
AC 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DR 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
GN (Hepatopoietin A).
OS Name=Hgf;
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshiro K., Hagiya M., Nishizawa T., Seki T., Shimomishi M.,
RA Shimizu S., Nakamura T.,
RT "Deduced primary structure of rat hepatocyte growth factor and
RT expression of the mRNA in rat tissues.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "Primary structure of rat hepatocyte growth factor and induction of
RT its mRNA during liver regeneration following hepatic injury.";
RL Eur. J. Biochem. 193:375-381(1990).
CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal
CC hepatocyte cells, seems to be an hepatotrophic factor, and acts as
CC growth factor for a broad spectrum of tissues and cell types. It
CC has no detectable protease activity.
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
CC disulfide bond.
CC -!- SIMILARITY: Belongs to the peptidase S1 family. Plasminogen
CC subfamily.
CC -!- SIMILARITY: Contains 4 kringle domains.
CC -!- SIMILARITY: Contains 1 PAN domain.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D90102; BA014133.1; -
CC EMBL: X54400; CA38256.1; -
CC PIR: A35644; A35644.
CC HSSP: P14210; 1BHT.
CC MEROPS: S01.978; -.
CC RGD: 2794; HGF.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR003014; PAN.
CC InterPro: IPR003609; Pan_app.
CC InterPro: IPR009003; Pept_Ser_Cys.
CC InterPro: IPR001254; Peptidase_S1.
CC InterPro: IPR001314; Peptidase_S1A.
CC Pfam: PF00051; Kringle; 4.
CC Pfam: PF00024; PAN; 1.
CC Pfam: PF00089; Trypsin; 1.
CC PIRSF: PIRSF001152; HGF_MST1; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle; 4.
CC SMART: SM00130; KR; 4.
CC SMART: SM00473; PAN_AP; 1.
CC SMART: SM00020; TRYP_SPC; 1.
CC PROSITE: PS00021; KRINGLE_1; 4.
CC PROSITE: PS0070; KRINGLE_2; 4.
CC PROSITE: PS50948; PAN; 1.
CC PROSITE: PS50240; TRYPSIN_DOM; 1.
CC Kx Direct protein sequencing; Glycoprotein; Growth factor; Kringle;
CC Pyroldione carboxylic acid; Repeat; Serine protease homolog; Signal.
CC FT SIGNAL 1 32
CC FT CHAIN 33 495
CC FT CHAIN 496 728
CC FT MOD_RES 33 33
CC FT DOMAIN 38 124
CC FT DOMAIN 129 207
CC FT DOMAIN 212 289
CC FT DOMAIN 306 384
CC FT DOMAIN 392 470
CC FT DOMAIN 496 728
CC FT DISULFID 71 97
CC FT DISULFID 75 85
CC FT DISULFID 607 67
CC FT CARBOHYD 295 295
CC FT CARBOHYD 403 403
CC FT CARBOHYD 569 569
CC FT CARBOHYD 656 656
CC FT SEQUENCE 728 AA; 82905 MW; 3E0BF1F6ADCEDF6 CRC64;

Query Match 92.2%; Score 2401; DB 1; Length 728;
Best local Similarity 90.4%; Pred. No. 1.1e-171;
Matches 404; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
Qy 1 EKKRRNTLHEFKKSAKTLTKEDPLKIKTKKVNNSADECANRCIRNKGFTTCKAFVFDK 60
Db 33 OKRRNTLHEFKKSAKTLTKEDPLKIKTKKVNNSADECANRCIRNKGFTTCKAFVFDK 92
Qy 61 ARKQCLMPFPNMSGCVKKEFGHEPDLVKNKYIRNCIIIGKGRSYGTVSITKSGIKCOP 120
Db 93 SKRRCYWPFPNMSGCVKKGFGHEPDLVKNKYIRNCIIIGKGSYKGTVSITKSGIKCOP 152
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCESEVC 180
Db 153 WMSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCESEVC 212

```

QY	161	MTMCNESYXGLMDHNESSGLCOQRWBOHTQTHRHXFLPERYPDGFDDNNCRNPQGORPMC	240
Db	213	MTMCNESYXGPMDDHTESSGTCQWRDQOTPHRHFLEPERYPDGFDDNNCRNPQGORPMC	272
QY	241	YTLDPHTRWYCAIKTCADNTMNDTVPLETTETCIGOGGBGYGTNTIWNIGJPCRCWRMS	300
Db	273	YTLDPDTPEWYCAIKMCAHSAVNETDVPMBETTECIGOGGBGYGTNTIWNIGJPCQWRMS	332
QY	301	QYPHEHDMTPENPFKCDLLENNYCRNPDGSESPMPCFTTDPRIIRGVYCSQIPNCMMSGDGC	360
Db	333	QYPHEHDDITPENPFKCDLLENNYCRNPDGASPMPCFTTDPRIIRGVYCSQIPKCVSSSGDGC	392
QY	361	YRGNKKNYVGNLSQTSRSGLTCSWMDKNMEDLHRHIFWEPDASKLANENYCRNPDDDAHGPW	420
Db	393	YRGNKKNYVGNLSKTRSGLTCSWMDKNMEDLHRHIFWEPDASKLTKNYCRNPDDDAHGPW	452
QY	421	CYTGNPPLIMWYDCPISRCGDDTPTIV	447
Db	453	CYTGNPPLVWYDCPISRCEGDDTPTIV	479

RESULT 8

ID	090978	PRELIMINARY;	PRT;	726 AA.
AC	090978	Q90866;		
DT	01-JAN-1998	(TREMBLrel. 05, Created)		
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)		
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)		
DE	Hepatocyte growth factor /scatter factor.			
GN	Name=HGF/SF;			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archiosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
NCBI	TaxID=9031;			
XX				

RP SEQUENCE FROM N.A.
RX MEDLINE=96029010; Pubmed=7554499;
RA Thery C., Sharpe M.J., Bailey S.J., Stern C.D., Gherardi E.,
RT "Expression of HGF/SF, HGF1/MSP and C-met suggests new functions
RT during early chick development";
RL Dev. Genet. 17:90-101(1995).
CC 1- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; X84045; CA58864.1; -.
DR PIR; I51285; I51285.
DR HSSP; P14210; IGMN.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004285; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR Pfam; PF00051; Kringle_4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PIRSF; PIRSF001152; HGT_MST1; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SMO0130; KR; 4.
DR SMART; SMO0473; PAN_AP; 1.
DR SMART; SMO0020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50948; PAN; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR HYDROLase; Kringl; Protease; Serine protease.
SQ SEQUENCE 726 AA; 82665 MW; 5BD06CFB5C40B003 CRC64;

Query Match 78.8%; Score 2052; DB 2; Length 726;

Best Local Similarity 76.6%; Pred. No. 1.6e-145;

Matches 341; Conservative 49; Mismatches 53; Indels 2; Gaps 1;

QY 3 KRANTHEEFKSAKTTLLKIDPAKIKTKKVNADQACRCTRNGKLPFTCKAPVEDKAR 62

Db	30	XRBNPLHYYKKTGELMLKVKAKTLEBVKTKLNTTQOCARCRNKGSLSTCCAFADRYT	89
QY	63	KQCIWFPPNSMSSGVYKEFGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKQOPMS	122
Db	90	KRCHLSNSLNTNGYRKKKDHFADLFEEKDYARNCIIGKAGAYKGTISITKSGIOCOAWN	149
QY	123	SMIPIHEHFLPSYVYKXKLOENYCNPNPGEEGPWCFTSNPEVRYEVCIPQSEVECMT	182
Db	150	SMIPIHEHFLPSYVYKXKLOENYCNPNPGEEGPWCFTSNPEVRYEVCIPQSEVECMT	209
QY	183	CNGESYRSRLMDHTSESGKICQRMWDHOTPHRKFLPERYPDKGFDNNYCRNPDGQPRWCYT	242
Db	210	CNGESYRCPMDHTSESGKQGRMDLQRPHKHKRPREYPDKGFDNNYCRNPDGKLRPMWYCT	269
QY	243	LDPHTRMEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGTVNTINAGIPQQRWDQY	302
Db	270	LDPNTPMEFCAIKTCDDVGLNSTEVAELTTTCIOGGEGYRGTVNTINAGIPQQRWDQY	329
QY	303	PHEHMTPEENPFCKDLRENYCNPAPGSESPWCFETDPNIRVYCSQIPQCDSMHGDDCYR	362
Db	330	PHQNHITPEENPFCKDLRENYCNPAPGSESPWCFETDPNIRVYCSQIPQCDSMHGDDCYR	389
QY	363	GNKKNYMNLSQTSRGLTCSMWDKNMEDLHRR--LFWEDASKLNEYNCRNPDDDAHGPW	420
Db	390	GNKGSYMNLSNSTRIGLTCSWTWCKDNIEDLRRIHQIFREPDVSKLKNYCNRNPDDDPHGPW	449
QY	421	CYTGNPPLIPWYCPISRCBGDTTPT	445
Db	450	CYTDDPLIPWYCPISRCBGDTTPT	474

RESULT 5

ID	Q91402	PRELIMINARY;	PRT.	710 AA.
AC	Q91402;			
DT	01-NOV-1996	(TREMBLrel. 01. Created)		
DT	01-NOV-1996	(TREMBLrel. 01. Last sequence update)		
DT	01-MAR-2004	(TREMBLrel. 26. Last annotation update)		
DE	Hepatocyte growth factor.			
GN	Name=HGF;			
OS	Xenopus.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OX	NCBI_TaxID=8353;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISUE=Tailbud;			
RX	MEDLINE=96267690. PubMed=7748783; DOI=10.1016/0925-4773(94)00309-B;			
RA	Nakamura H., Tashiro K., Nakamura T., ShioKawa K.;			
RT	"Molecular cloning of Xenopus HGF cDNA and its expression studies in			
RT	Xenopus early embryogenesis."			
RU	Mech. Dev. 49:123-131(1995).			
CC	-1. SIMILARITY: Belongs to peptidase family S1.			
DR	EMBL; S77422; AAB34354.2; -.			
DR	HSSP; P14210; IGMN.			
DR	MEROPS; S01_976; -.			
DR	GO: GO:0004263; F:chymotrypsin activity; IEA.			
DR	GO: GO:0008233; F:peptidase activity; IEA.			
DR	GO: GO:0004295; F:trypsin activity; IEA.			
DR	GO: GO:0006508; P:proteinolysis and peptidolysis; IEA.			
DR	Pfam; PF00051; Kringle_4.			
DR	Pfam; PF00024; PAN; 1.			
DR	Pfam; PF00089; Trypsin; 1.			
DR	PIRSF; PIRSF001152; HGF_MST1; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00018; KRINGLE.			
DR	Prodom; PD000395; Kringle_4.			
DR	SMART; SM00130; KR; 4.			
DR	SMART; SM00473; PAN_AP; 1.			
DR	SMART; SM00020; Tryp_Spc; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 4.			
DR	PROSITE; PS00070; KRINGLE_2; 4.			


```

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen.
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RC Klein S., Strausberg R.;
RA Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL, BC073334; AAH73334.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00024; Kringl; 4.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringl; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50948; PAN; 1.
KW Kringle.
SQ SEQUENCE 449 AA; 52467 MW; 5C2CE4B708E13D0B CRC64;

Query Match 46.6%; Score 1214; DB 2; Length 449;
Best Local Similarity 47.3%; Pred. No. 6; 8e-83;
Matches 207; Conservative 74; Mismatches 145; Indels 12; Gaps 6;

QY 5 RNTIHEFKKSAATTLIKIDPALKIKTKKYNVTADQCANRCTRNKGLPFTCKAFVDPKARQ 64
DB 21 RSLNDYQSKGLELVHMNNG-GVKQEIQSEIQVCAKQCS-----LLDCRSFYVMKWSQS 75

QY 65 CLMFPFNSSGVKKEFGHEFDLYENKDYINRCIIIGKSGSYGTVSITSSGIKQCPMSM 124
DB 76 CLLPFWTONSAVLLQQRNVQYDLQYKQDYIRCCVWNGNTYTGYSKTSKSGTCCORWRLK 135

QY 125 IPHEHSFLPSSYRGKDLQENYCRNPGEEGGPMCFSTNSPEYRYEVDIPQCESEVEMCTN 184
DB 136 FPHDHKFSPIHW--PULEENYCRNPDSDEEGPMCYTTDKNIRHQYIGIKCEDAVCLTN 193

QY 165 GESYGLMDHTESGKICQRMWHQTPHRAKFLPERYPDKGFDNDYCRNPDGQRPWCYTLD 244
DB 194 GSDYGSVDRTEGSKGECQRMWLOTPLAHAPYKPEKYPDKSLDNDYCRNPDSSEPMCYTTD 253

QY 245 PHTREHYCAIKTCADNTMNDTVPLETTETCIQGGSGYGTWNTWNGIPCCQRMWSQYH 304
DB 254 PAVEKEFCITKCKQKRLSNIET--TSYCFKERGGYGGKANTTSSGIPCCQRMWSQYH 310

QY 305 EHDMTPENFCKDLRENYCRNPDSGSPMCFSTTDPNIRVGYCSQIPNC-DMSHGQDCYRG 363
DB 311 SHRFLPEKTPCKGLDENYCRNPDSGSPMCFSTTLPQMRMAVCFQIKRCQDVLLEPDCYHG 370

```

```

QY 364 NGKMYGNLSQTRSGLTCSMDKNNEDLHRIHFWEPDASKLNENYCRNPDDAHGPMCYT 423
DB 371 NGELISGVRGSTRKIKRCRWEKRNLDLESLD-QPYLVPLEENYCRNPDRDSHGPMCYT 429

QY 424 GNPLIPMDYCPISRCESD 441
DB 430 MDNTPFDYCAIKPCEGE 447

RESULT 12
Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN Name=HGF/MSP;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF/MSP and c-met suggests new functions
RT during early chick development."
RL Dev. Genet. 17:90-101(1995).
CC -i SIMILARITY: Belongs to peptidase family S1.
DR EMBL, X80433; CAA58862.1; -.
DR HSSP; P00747; ICEA.
DR MEROPS; S01.977; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR Pfam; PF00051; Kringl; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PIRSF; PIRSF001152; HGF_MST1; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringl; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spe; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50948; PAN; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Hydrolase; Kringl; Protease; Serine protease.
SQ SEQUENCE 704 AA; 79341 MW; CAB0D8C41367C37 CRC64;

Query Match 46.4%; Score 1209.5; DB 2; Length 704;
Best Local Similarity 46.0%; Pred. No. 2; 4e-82;
Matches 207; Conservative 71; Mismatches 149; Indels 23; Gaps 8;

QY 5 RNTIHEFKKSAATTLIKIDPALKIKTKKYNVTADQCANRCTRNKGLPFTCKAFVDPKARQ 64
DB 20 RSLNDFORLGTETELRAAPNEPPSPAPAHGAQCCACQCANRP-----DCRAFHHEROSL 75

QY 65 CLMFPFNSSGVKKEFGHEFDLYENKDYINRCIIIGKSGSYGTVSITSSGIKQCPMSM 124
DB 76 COLLPWSQSPSPARLQKNIHYDLQYKQDFLBCIVANGSTRGTRDITTEGRGLRCHWQAT 135

QY 125 IPHEHSFLPSSYRGKDLQENYCRNPGEEGGPMCFSTNSPEYRYEVDIPQCESEVEMCTN 184
DB 136 TPHDHRLFLPSLRNG--LEENYCRNPDRDKRGPMCYTTVDPNVRHQSGGIIKCEDAVCMCTN 193

QY 165 GESYGLMDHTESGKICQRMWHQTPHRAKFLPERYPDKGFDNDYCRNPDGQRPWCYTLD 244

```


Db 194 GEYRFEVHTESGTECQRMWDLQHPKHHPDPKYEKGJLDNVCNPPSSSEPCWCTYTD 253
QY 245 PHTREYCAIKTCADNTMNDVPLR-ETTECIGOGEGYRGVNTIWMGIPCORWDSOYP 303
Db 254 PALERFECIRYVKRPR-----PIVTTTCIRGKEGEGIRGRNVTVSGIPCORWDAQTL 308
QY 304 HEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDNIRVYCSQIPNCDSM-HGQDCYR 362
Db 309 HRHFPVSKYPCKDLOENVCNPDGSEBAPWCFTTRGMRVAFCEHTRCDDDELDAECVH 368
QY 363 GNGKATMGSLQTRSGLTCSMDKMDLHRH1-----FWEPDASKLNNVCNPDDDAH 417
Db 369 GHEERYHGHVSKTRKGIITCORWDATTP---HYPOISPTTHPEA-HLEKNVCNPPDSDH 423
QY 418 GPCWCTGNPLIPWDYCPISRCBGDTPTIY 447
Db 424 GPCWCTGNPLIPWDYCPISRCBGDTPTIY 453

RESULT 13

091691 PRELIMINARY; PRT; 716 AA.
AC 091691;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE Growth factor Liverline.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ruiz i Altaba A., Thery C.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U57455; AAB52574.1; -
DR HSSP; P00746; 1FDP.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:fibrinogen activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PIRSF; PIRSF001152; HGF_MST1; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_A1; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00948; PAN; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR HydroLase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match

Best Local Similarity 46.4%; Score 1208; DB 2; Length 716;
Matches 207; Conservative 75; Mismatches 149; Indels 12; Gaps 6;

QY 5 RNTHEFFKSAKTLIKIPALIKITKKNVTAQOCANRCTRNKGLPFTCAEFVFKARQ 64
Db 31 RSLANDYGRSKGLIELVHMANG-GVKEIOSEIQVCAKQSD-----LLDGRSFYVNMKSGT 85
QY 65 CLWFPNMSGGVKGFEHGFLENDKYIKNCTICKGRSYKGTVSYTSGIKCOPWSSM 124

Db 86 CELLFWTONSANVLLQRRNVQYDLQYKQDYIRPCVANGNTYRGTVSKTSKGRTCQRMRLK 145
QY 125 IHEHSFLPSSYRGKDLQENVCNPDGSESPWCFTTDNIRVYCSQIPNCDSM-HGQDCYR 184
Db 146 FPHDKFSPFHW--BELENNVCNPDSDPEGWCFTTDNIRVYCSQIPNCDSM-HGQDCYR 203
QY 185 GSYGLMDHTESGKICQRMHQTQHRHKLPERYPDKGFDNVCNPDGPPWCTYTD 244
Db 204 GSDYVGSVDRTESGKECQRMWDLQTPHAPYKREKPDKSLDNNVCNPPSSSEPCWCTYTD 263
QY 245 PHTREYCAIKTCADNTMNDVPLR-ETTECIGOGEGYRGVNTIWMGIPCORWDSOYP 304
Db 264 PVEKEBFRCITCKKQRPISNIEI---TSTCFKEREGERGKANTTSGIPCORWDSOYP 320
QY 305 HEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDNIRVYCSQIPNCDSM-HGQDCYR 363
Db 321 SHRPPEKYPCKGLDENVCNPDGSEBAPWCFTTRGMRVAFCEHTRCDDDELDAECVH 380
QY 364 GNGKATMGSLQTRSGLTCSMDKMDLHRH1FWEPDASKLNNVCNPDDDAHGPCWCT 423
Db 381 NGELYSGRVSKTRKGIKCRMEKEKNDLESLD-QYLVPLLENNVCNPDSDHGPCWCT 439
QY 424 GNPILIPWDYCPISRCBGDTPTIY 446
Db 440 MDPNTPFPDYCAIKPCEGEKVTL 462

RESULT 14

HGFL_HUMAN STANDARD; PRT; 711 AA.
ID HGFL_HUMAN
AC P26927; Q13350; Q14870;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, last sequence update)
DT 25-OCT-2004 (Rel. 45, last annotation update)
DE Hepatocyte growth factor (MSP) (Macrophage stimulatory protein).
GN Name=MST1; Synonyms=HGF,
Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92002016; PubMed=1655021;
RA Han S., Stuart L.A., Friesner Degen S.J.;
RT "Characterization of the DNF152 locus on human chromosome 3;
RT Identification of a gene coding for four kringle domains with homology
RT to hepatocyte growth factor.";
RN Biochemistry 30:9768-9780(1991).
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93340141; PubMed=8393443;
RA Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
RT "Cloning, sequencing, and expression of human macrophage stimulating
RT protein (MSP, MST1) confirms MSP as a member of the family of kringle
RT proteins and locates the MSP gene on chromosome 3.";
RN J. Biol. Chem. 268:15461-15468(1993).
RN [3]
RP SEQUENCE OF 230-247; 288-310; 326-341; 484-501; 530-549; 574-596 AND
RP 602-611, AND SUBUNIT.
RC TISSUE=Plasma;
RX PubMed=1827141;
RA Skeel A., Yoshimura T., Showalter S.D., Tanaka S., Appella E.,
RT "Macrophage stimulating protein: purification, partial amino acid
RT sequence, and cellular activity".
J. Exp. Med. 173:1227-1234(1991).
CC FUNCTION: Probably has no proteolytic activity, since crucial AA
CC characteristic of serine proteases catalytic sites are not
CC conserved.

CC -1- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
 CC disulfide bond.
 CC -1- PTM: May be cleaved after Arg-483, to yield two chains held
 CC together by disulfide bonds, or two separate polypeptides.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family. Plasminogen
 CC subfamily.
 CC -1- SIMILARITY: Contains 4 kringle domains.
 CC -1- SIMILARITY: Contains 1 PAN domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M74178; AA50165.1; -
 CC EMBL, U37055; AAC50471.1; -
 CC EMBL, L11924; AA59872.1; -
 CC PIR, A40331; A47136.
 CC HSSP, P00766; 1CHG.
 CC MEROPS, S01.975; -
 CC Genew; HNC:7380; MST1.
 CC MIM, 142408; -
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR003014; PAN.
 CC InterPro; IPR003609; Pan app.
 CC InterPro; IPR009003; Pept_Ser_Cys.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00051; Kringle; 4.
 CC Pfam; PF00024; PAN; 1.
 CC Pfam; PF00089; Trypsin; 1.
 CC PIRSF, PIRSF001152; HGF_MST1; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC PRODOM; PD000395; Kringle; 4.
 CC SMART; SM00130; KR; 4.
 CC SMART; SM00473; PAN_AP; 1.
 CC SMART; SM00020; TRYP_SPC; 1.
 CC PROSITE; PS00021; KRINGLE_1; 4.
 CC PROSITE; PS50048; KRINGLE_2; 4.
 CC PROSITE; PS50949; PAN; 1.
 CC PROSITE; PS50240; TRYPSIN_DOM; 1.
 CC Direct protein sequencing; Glycoprotein; Kringle; Polymorphism;
 CC Repeat; Serine protease homolog; Signal.
 CC SIGNAL
 CC CHAIN 1 19 711 18
 CC FT DOMAIN 1 21 105 186
 CC FT DOMAIN 110 186
 CC FT DOMAIN 191 268
 CC FT DOMAIN 283 361
 CC FT DOMAIN 370 448
 CC FT DOMAIN 484 711
 CC FT DISULFID 56 78
 CC FT DISULFID 60 66
 CC FT DISULFID 110 186
 CC FT DISULFID 131 169
 CC FT DISULFID 157 181
 CC FT DISULFID 191 268
 CC FT DISULFID 194 324
 CC FT DISULFID 212 251
 CC FT DISULFID 240 263
 CC FT DISULFID 283 361
 CC FT DISULFID 304 343
 CC FT DISULFID 332 355
 CC FT DISULFID 370 448
 CC FT DISULFID 391 431
 CC FT DISULFID 419 443
 CC FT DISULFID 468 588

FT DISULFID 507 523 By similarity.
 FT DISULFID 602 667 By similarity.
 FT DISULFID 632 666 By similarity.
 FT DISULFID 657 685 By similarity.
 FT CARBOHYD 72 72 N-linked (GlcNAc...) (Potential)
 FT CARBOHYD 296 296 N-linked (GlcNAc...) (Potential)
 FT CARBOHYD 615 615 N-linked (GlcNAc...) (Potential)
 FT VARIANT 13 13 Y->C.
 FT VARIANT 212 212 C->F.
 FT VARIANT 676 676 E->K (in dbSNP:7798).
 FT VARIANT 676 676 /FTID=VAR_014569.
 FT CONFLICT 292 292 R->G (in Ref. 3).
 FT CONFLICT 304 304 C->E (in Ref. 3).
 FT CONFLICT 306 306 R->E (in Ref. 3).
 FT CONFLICT 550 551 PS -> ST (in Ref. 3).
 FT CONFLICT 593 593 W->E (in Ref. 3).
 FT CONFLICT 623 623 L->F (in Ref. 2).
 SQ SEQUENCE 711 AA; 80379 MW; 596ED21F180290E4 CRC64;

Query Match 45.3%; Score 1180.5; DB 1; Length 711;
 Beet Local Similarity 47.4%; Pred. No. 3.7e-80; Indels 9; Gaps 4;
 Matches 197; Conservative 62; Mismatches 146;

QY 33 VNTADQCANRCTRNRKGLPFTCKAFVFDKARKQCLMFPPNSMSGVKEFGHEFDLYENKD 92
 DB 50 VADAECAGRG---GPLMDGRAFHVNVSHGCCQLPWTQSHPHRLRSGRCDFQKXD 105
 QY 93 YIRNCIICKGKSYKTVSITSGIKCQPMSSMIPHEHSFLLSSIRGKLOENYCRNPGE 152
 DB 106 YVRCITMNVGVGYRVTATVGGVPCQAMSHKFPDHRXYPTLNLNG--LEENFCNPGD 163
 QY 153 EGGPWCFTSNPEVREVDIPQCSVEBCMTNGSGSYRGLMDHTSGKICQRMQDTPRH 212
 DB 164 PGGPWCYTTPDAVRFQSGICSCREACVWNGEYKGAUVRTSGRCQMDLQHPQH 223
 QY 213 KFLPERYDKFPDDNYCRNPDGQPPRPMWCYTLDPHTRWEYCAIKTCADNTMNDTVELETT 272
 DB 224 PFEPEKFLDQGLDDVYCNPDGSERPWCYTLDQLERFCDLPCGSAQROEA--TTV 281
 QY 273 ECIQCGSGRYGTVTNTNGIPQQRWDSQYRHEHDMTEFNKCDLRENYCRNPGSSSP 332
 DB 282 SCFRKGKGYRGYATATTAGVPCQRMDOIPHQHFTEKACDLENFRCNPDGSAAP 341
 QY 333 WCFTDPNIRVGYCQIPNC-DMSHGDCYRGNGKNYNGNLSQTRSGLTCSMMDKNMEDL 391
 DB 342 WCFILRPGRARAFCTQIRCTDDVRFQDCTYGAAGEQYRGYVSKTRKGYQCQRMASRTHK 401
 QY 392 HRHIFMEBDASKLNNYCRNPDGDAHGPWCYTGNPLIPWDYCPISRCGDTPTTV 447
 DB 402 PQFTTSPHQLLEENFCRNPFGDSHGPMCYTMDPRTFDFYCALRCADDQPPSL 457

RESULT 15
 P70006 PRELIMINARY; PRT; 717 AA.
 AC P70006; 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney and Liver;
 RX MEDLINE=96404125; PubMed=8808403; DOI=10.1016/0925-4773(95)00458-0;
 RA Aberger F., Schmidt G., Richter K.;
 "The Xenopus homolog of hepatocyte growth factor-like protein is

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:29 ; Search time 172.465 Seconds
(without alignments)
1002.420 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604
Sequence: 1 ERKRRNTIHEFKKSAKTLLI.....IPWDYCPISRCGDTPTTIV 447

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2601	99.9	447	7	ADE37393	Ade37393 Human hep
2	2601	99.9	447	7	ADM83266	Adm83266 Human NK4
3	2601	99.9	697	3	AAV98485	Aav98485 Hepatocyt
4	2601	99.9	697	3	AAV59030	Aav59030 Sequence
5	2601	99.9	697	4	AAAB5838	Aab5838 Nucleic a
6	2601	99.9	697	4	AAU04275	Aau04275 Nuclear 1
7	2601	99.9	728	2	AAAR25676	Aar25676 Recombina
8	2601	99.9	728	2	AAAR20005	Aar20005 Human hep
9	2601	99.9	728	2	AAAR40862	Aar40862 Competati
10	2601	99.9	728	2	AAAR40863	Aar40863 Competati
11	2601	99.9	728	2	AAAR42062	Aar42062 Vascular
12	2601	99.9	728	2	AAAR87522	Aar87522 Mutant he
13	2601	99.9	728	2	AAAR87524	Aar87524 Mutant he
14	2601	99.9	728	2	AAAR87525	Aar87525 Mutant he
15	2601	99.9	728	2	AAAR87523	Aar87523 Mutant he
16	2601	99.9	728	2	AAAW00340	Aaw00340 Wild type
17	2601	99.9	728	2	AAAW00338	Aaw00338 Human hep
18	2601	99.9	728	2	AAAW59922	Aaw59922 Human leu
19	2601	99.9	728	2	AAAW58696	Aaw58696 Human hep
20	2601	99.9	728	2	AAAW42998	Aaw42998 Recombina
21	2601	99.9	728	2	AAAW39207	Aaw39207 Human hep
22	2601	99.9	728	6	ABPS8128	Abp8128 Human hep
23	2601	99.9	728	7	ADB61534	Adb61534 Hepatocyt
24	2601	99.9	728	8	ADM97652	Adm97652 Human hep
25	2601	99.9	728	8	ADT89322	Adt89322 Human hep

26	2601	99.9	951	8	ADSI9044	Adsi9044 Chimeric
27	2599	99.8	447	3	AAV57173	Aav57173 N-termi
28	2595	99.7	727	2	AAAI0656	Aai0656 Hepatic p
29	2595	99.7	728	2	AAAI5623	Aai5623 Human leu
30	2595	99.7	728	2	AAAI44243	Aai44243 Human hep
31	2595	99.7	728	2	AAI39521	Aai39521 Hepatocyt
32	2591	99.5	728	2	AAW88529	Aaw88529 Human hep
33	2588	99.4	728	2	AAW88532	Aaw88532 Human hep
34	2587	99.3	728	2	AAW88531	Aaw88531 Human hep
35	2586	99.3	728	2	AAW88530	Aaw88530 Human hep
36	2577	99.0	728	2	AAAR47227	Aar47227 Hepatocyt
37	2565	98.5	728	2	AAAR52942	Aar52942 Human hep
38	2565	98.5	728	2	AAAR52941	Aar52941 Human hep
39	2565	98.5	728	2	AAAR52942	Aar52942 Human hep
40	2565	98.5	728	2	AAAR52945	Aar52945 Human hep
41	2565	98.5	728	2	AAAR52947	Aar52947 Human hep
42	2565	98.5	728	2	AAAR52944	Aar52944 Human hep
43	2565	98.5	728	2	AAAR52948	Aar52948 Human hep
44	2565	98.5	728	2	AAAR52940	Aar52940 Human hep
45	2565	98.5	728	2	AAAR52943	Aar52943 Human hep

ALIGNMENTS

RESULT 1	ADBE37393	ADBE37393 standard; protein; 447 AA.
ID	ADBE37393	
XX	ADBE37393;	
XX	29-JAN-2004 (first entry)	
DT	29-JAN-2004 (first entry)	
XX	Human hepatocyte growth factor N-terminal four kringle fragment.	
DE	Human hepatocyte growth factor N-terminal four kringle fragment.	
XX	conjugate; HGF/SF; hairpin domain; kringle region; PBG group; cancer;	
KW	human;	
KW	hepatocyte growth factor N-terminal four kringle-containing fragment;	
KM	NK4.	
XX		
OS	Homo sapiens.	
XX		
PN	EP1234583-A1.	
XX		
PD	28-AUG-2002.	
XX		
PF	23-FEB-2001; 2001EP-00104640.	
XX		
PR	23-FEB-2001; 2001EP-00104640.	
XX		
PA	(HOPE) HOFFMANN LA ROCHE & CO AG F.	
XX		
DR	WPI, 2003-749561/71.	
XX		
PT	N-PSDB; ADE37392.	
XX		
PT	Conjugate comprises an N-terminal fragment of hepatocyte growth factor	
PT	consisting of the hairpin domain and the four kringle regions of the	
PT	alpha-chain and one to three polyethylene glycol groups, useful for	
PT	treating cancer.	
XX		
PS	Disclosure; SEQ ID NO 2; 25pp; English.	
XX		
CC	The invention comprises a conjugate that contains an N-terminal fragment	
CC	of hepatocyte growth factor (HGF/SF) consisting of the hairpin domain and	
CC	the four kringle regions of the alpha-chain and 1-3 polyethylene glycol	
CC	(PEG) groups. The conjugate of the invention is useful for the	
CC	preparation of a medicament for the treatment of cancer. The present	
CC	amino acid sequence represents the human N-terminal four kringle-	
XX	containing fragment of hepatocyte growth factor (NK4).	
SO	Sequence 447 AA;	
Query Match	99.9%; Score 2601; DB 7; Length 447;	

Best Local Similarity 99.8%; Pred. No. 5.5e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRNTIHEFKSKAKTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCAFAVFDK 60
DB 1 QKRRNTIHEFKSKAKTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCAFAVFDK 60
QY 61 ARKQCLMPFMSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKCP 120
DB 61 ARKQCLMPFMSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKCP 120
QY 121 WSSMTPIHSHSLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVFC 180
DB 121 WSSMTPIHSHSLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVFC 180
QY 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHHRKFLPERYPDKGPDNYCRNPDGQPRWC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHHRKFLPERYPDKGPDNYCRNPDGQPRWC 240
QY 241 YTLDPHTMEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGVTIWMGIPCORWDS 300
DB 241 YTLDPHTMEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGVTIWMGIPCORWDS 300
QY 301 QYPHEHDMTPENPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 301 QYPHEHDMTPENPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
QY 361 YRNGKNYMGNLSTQRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 420
DB 361 YRNGKNYMGNLSTQRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 420
QY 421 CYTGNPLIPMDYCPISRCGSDTTPTIIV 447
DB 421 CYTGNPLIPMDYCPISRCGSDTTPTIIV 447

RESULT 2
ADM83266
ID ADM83266 standard; protein; 447 AA.
AC ADM83266;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human NK4 protein #1.
XX
KW NK4 gene; neovascularisation; ovarian cancer; pancreatic cancer;
KW stomach cancer; gallbladder cancer; renal cancer; prostatic cancer;
KW breast cancer; esophageal cancer; hepatic cancer; oral cavity cancer;
KW colon cancer; colorectal cancer; sarcoma; glioma; melanoma; gene therapy;
KW human.
XX
OS Homo sapiens.
XX
PN US2003162736-A1.
XX
PD 28-AUG-2003.
XX
PE 27-SEP-2002; 2002US-00255649.
XX
PR 25-FEB-2002; 2002JP-00048644.
XX
PA (NAKA/) NAKAMURA T.
XX (MATS/) MATSUMOTO K.
XX
PI Nakamura T, Matsumoto K;
XX
XX WPI; 2003-897946/82.
XX DR N-PSDB; ADM83264.
XX
XX
PT New NK4 gene, useful for producing a medicament for preventing or
XX treating diseases caused by neovascularization, e.g. cancer.

PS Disclosure; Page 15-17; 27pp; English.
XX
XX The present invention provides a therapeutic agent NK4 polypeptides and
CC their encoding polynucleotides. The invention is useful for producing a
CC medicament for preventing or treating diseases caused by
CC neovascularisation such as ovarian cancer, pancreatic cancer, stomach
CC cancer, gallbladder cancer, renal cancer, prostatic cancer, breast
CC cancer, esophageal cancer, hepatic cancer, oral cavity cancer, colon
CC cancer, colorectal cancer, sarcoma, glioma and melanoma. The invention is
CC also useful in gene therapy. The present sequence is human NK4 protein.
XX
SQ Sequence 447 AA;
Query Match 99.9%; Score 2601; DB 7; Length 447;
Best Local Similarity 99.8%; Pred. No. 5.5e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRNTIHEFKSKAKTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCAFAVFDK 60
DB 1 QKRRNTIHEFKSKAKTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCAFAVFDK 60
QY 61 ARKQCLMPFMSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKCP 120
DB 61 ARKQCLMPFMSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKCP 120
QY 121 WSSMTPIHSHSLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVFC 180
DB 121 WSSMTPIHSHSLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVFC 180
QY 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHHRKFLPERYPDKGPDNYCRNPDGQPRWC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHHRKFLPERYPDKGPDNYCRNPDGQPRWC 240
QY 241 YTLDPHTMEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGVTIWMGIPCORWDS 300
DB 241 YTLDPHTMEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGVTIWMGIPCORWDS 300
QY 301 QYPHEHDMTPENPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 301 QYPHEHDMTPENPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
QY 361 YRNGKNYMGNLSTQRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 420
DB 361 YRNGKNYMGNLSTQRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 420
QY 421 CYTGNPLIPMDYCPISRCGSDTTPTIIV 447
DB 421 CYTGNPLIPMDYCPISRCGSDTTPTIIV 447

RESULT 3
AAV98485
ID AAV98485 standard; protein; 697 AA.
XX
XX AAV98485;
XX
DT 31-JUL-2000 (first entry)
XX
DE Hepatocyte growth factor Peg 20 used in nucleic acid transporter system.
XX
KW Transporter system; nucleic acid delivery; gene therapy; cancer;
KW carcinogenesis; cardiovascular disease; infection.
XX
OS Synthetic.
XX
PN US6033884-A.
XX
PD 07-MAR-2000.
XX
PE 14-DEC-1993; 93US-00167641.
XX
XX 20-MAR-1992; 92US-00855389.
XX PR 19-MAR-1993; 93WO-US002725.

XX (BAYU) BAYLOR COLLEGE MEDICINE.
XX
XX Gotchalk S, Sparrow J, Cristiano RJ, Woo SLC, Smith LC;
XX WPI; 2000-281993/24.
XX
XX System for transporting nucleic acid into cells, useful e.g. in gene
PT therapy and for generating transgenic animals, comprises binding agent
PT linked to nucleic acid, surface ligand and lytic agent.
XX
XX
PS Disclosure; Fig 23A; 108bp; English.
XX
XX The present invention relates to a transporter system for delivering
CC nucleic acid to a cell. The system comprises a nucleic acid binding
CC complex, consisting of a binding molecule bonded non-covalently to the
CC nucleic acid, and covalently to a surface ligand, and a lytic agent. The
CC binding molecule is spermine or a spermidine derivative. Nucleotide
CC sequences AAA36633-A36652 and peptide sequences AAY9456-Y98500 are used
CC in the construction of the transporter system of the invention. The
CC transporter system is used in gene therapy, particularly to deliver
CC nucleic acids to hepatocytes, muscle cells or bone forming cells, e.g. for
CC treating cardiovascular disease, cancer, and infection. The transporter
CC systems are also used to create transgenic animals (as models for human
CC carcinogenesis or disease or for drug testing). Other uses include
CC transforming cells to produce proteins, or transfecting cells in vitro
CC to study the function of the nucleic acid. The use of a surface ligand
CC allows specific targeting of selected cells and tissues. The lytic agent
CC provides for release of the nucleic acid into the cellular interior, from
CC endosomes, without requiring endosomal or lysosomal degradation
XX
XX Sequence 697 AA;
SQ
Query Match 99.9%; Score 2601; DB 3; Length 697;
Best Local Similarity 99.8%; Pred. No. 8.7e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERRRRNTIEFKKSATTLTIKIDPALKIKTKKVNNTADOCANRCTRNGKLPTCKAFVFDK 60
DB 1 QRRRRNTIEFKKSATTLTIKIDPALKIKTKKVNNTADOCANRCTRNGKLPTCKAFVFDK 60
QY 61 ARKQCLMPFPNSMSSGVKKEFGHEPDLYENKQYIRNCIIIGKRSYKGTVSITSGIKCOP 120
DB 61 ARKQCLMPFPNSMSSGVKKEFGHEPDLYENKQYIRNCIIIGKRSYKGTVSITSGIKCOP 120
QY 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVYEVCDIPQCSVEVC 180
DB 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVYEVCDIPQCSVEVC 180
QY 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPDGQPRPWC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPDGQPRPWC 240
QY 241 YTLDPHTRMWYCAIKTCADNTMNDTVPLETTECICQCGSGYGTVNTTNGIPQCRWDS 300
DB 241 YTLDPHTRMWYCAIKTCADNTMNDTVPLETTECICQCGSGYGTVNTTNGIPQCRWDS 300
QY 301 QYHEHDMTPENKCKDLRENYCRNPDGSGSPWCFTTDPNIRVGYCSQILPNCMSHGQDC 360
DB 301 QYHEHDMTPENKCKDLRENYCRNPDGSGSPWCFTTDPNIRVGYCSQILPNCMSHGQDC 360
QY 361 YRNGKXNYMGNLSQTSRSGILCSMWDKXNMDLHRIHFWEPDASKLNNYCRNPDDAHGPV 420
DB 361 YRNGKXNYMGNLSQTSRSGILCSMWDKXNMDLHRIHFWEPDASKLNNYCRNPDDAHGPV 420
QY 421 CYTGNPLIPWDYCPISRCBGDTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCBGDTPTIV 447
RESULT 4
AAY59030
ID AAY59030 standard; peptide; 697 AA.

XX
XX AAY59030;
XX
XX 07-MAR-2000 (first entry)
XX
XX Sequence of a peptide ligand Pgp20.
XX
XX Nucleic acid transport system; NTS; cell surface receptor; cytosols;
XX nuclear membrane; lysis moiety; transgenic animal; human disease;
XX nucleic acid delivery; cancer.
XX
XX Synthetic.
XX
XX US5994109-A.
XX
XX 30-NOV-1999.
XX
XX 03-JUN-1995; 95US-00460890.
XX
XX 20-MAR-1992; 92US-00855389.
XX 19-MAR-1993; 93MO-US002725.
XX 14-DEC-1993; 93US-00167641.
XX
XX (BAYU) BAYLOR COLLEGE MEDICINE.
XX
XX Woo SLC, Cristiano RJ, Gotchalk S, Sparrow J, Smith LC;
XX WPI; 2000-038262/03.
XX
XX Nucleic acid transport system, useful for creating transgenic animals for
PT assessing human disease such as cancer in an animal model.
XX
XX
XX Disclosure; Fig 23A; 107bp; English.
XX
XX The invention relates to a nucleic acid transport system (NTS) for
CC delivering nucleic acid into a cell. The NTS contains but is not limited
CC to 5 components: (a) the nucleic acid or a macromolecule to be delivered;
CC (b) a moiety that recognizes and binds to a cell surface receptor or
CC antigen or is capable of entering a cell through cytosols; (c) a nucleic
CC acid or macromolecular molecule binding moiety; (d) a moiety that is
CC capable of moving or initiating movement through a nuclear membrane; and/
CC or (e) a lysis moiety that enables the transport of the entire complex
CC from the cell surface directly into the cytoplasm of the cell. The NTS
CC delivers nucleic acid into the cellular interior as well as the nucleus
CC of specific cells. The NTS can be used to treat disorders by targeting
CC specific nucleic acid accordingly. The NTS can also be used to create
CC transgenic animals for assessing human disease, such as cancer, in an
CC animal model. The NTS can be used in vitro with tissue culture cells
CC which allows the role of various nucleic acids to be studied by targeting
CC specific expression into specifically targeted tissue culture cells. The
CC lysis agent within the NTS avoids the problem of endosomal/lysosomal
CC degradation
XX
XX Sequence 697 AA;
SQ
Query Match 99.9%; Score 2601; DB 3; Length 697;
Best Local Similarity 99.8%; Pred. No. 8.7e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERRRRNTIEFKKSATTLTIKIDPALKIKTKKVNNTADOCANRCTRNGKLPTCKAFVFDK 60
DB 1 QRRRRNTIEFKKSATTLTIKIDPALKIKTKKVNNTADOCANRCTRNGKLPTCKAFVFDK 60
QY 61 ARKQCLMPFPNSMSSGVKKEFGHEPDLYENKQYIRNCIIIGKRSYKGTVSITSGIKCOP 120
DB 61 ARKQCLMPFPNSMSSGVKKEFGHEPDLYENKQYIRNCIIIGKRSYKGTVSITSGIKCOP 120
QY 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVYEVCDIPQCSVEVC 180
DB 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVYEVCDIPQCSVEVC 180
QY 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPDGQPRPWC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPDGQPRPWC 240

Db 181 MTCNGESYRGIMDHTESGKICQRMHDQTPHHRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240
Qy 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGVTNIMGIFCQWRDS 300
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGVTNIMGIFCQWRDS 300
Qy 301 QYPHEHMTPEPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 301 QYPHEHMTPEPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
Qy 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420
Db 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420
Qy 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

RESULT 5

AAB45838
ID AAB45838 standard; protein; 697 AA.

XX AAB45838;

XX 21-MAR-2001 (first entry)

XX Nucleic acid transporter system peptide ligand SEQ ID NO 50.

XX Nucleic acid delivery; nucleic acid transporter system; hormone; enzyme;
KW growth factor; clotting factor; apolipoprotein; receptor; drug; oncogene;
KW tumor antigen; tumor suppressor; viral antigen; parasitic antigen;
KW bacterial antigen.

XX Unidentified.

XX US6150168-A.

XX 21-NOV-2000.

XX 05-JUN-1995; 95US-00460971.

XX 20-MAR-1992; 92US-00855389.

XX 19-MAR-1993; 93WO-US002725.

XX 14-DEC-1993; 93US-00167641.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Gottchalk S, Sparrow J, Cristiano RJ, Smith LC, Woo SLC;

XX WPI; 2001-049093/06.

XX Nucleic acid transporter system for delivering nucleic acid into a cell,
PT useful for delivering proteins and polypeptides to cells, including
PT growth factors, enzymes, hormones, and tumor suppressors.

XX Claim 8; Col 115-118; 105pp; English.

XX This invention describes a novel system (I) for delivering a nucleic acid
CC to a cell, comprising a binding complex comprising a ligand binding
CC molecule noncovalently bound to a nucleic acid and covalently linked to a
CC surface ligand, and a second binding complex comprising a second binding
CC molecule noncovalently bound to a nucleic acid and covalently linked to a
CC nuclear ligand. The complexes are simultaneously bound to the nucleic
CC acid. The nucleic acid transporter system can also be used in a method
CC for the in vivo targeting of the insertion of DNA into a cell. It can
CC also be used in processes for producing transformed cell lines. The
CC system can be used to deliver a variety of proteins and polypeptides,
CC such as hormones, growth factors, enzymes, clotting factors,
CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor
CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.
CC The transporter system uses lysis agents to overcome the problems of
CC endosomal/lysosomal degradation seen with prior art systems

XX
SQ Sequence 697 AA;

Query Match 99.9%; Score 2601; DB 4; Length 697;
Best Local Similarity 99.8%; Pred. No. 8.7e-155;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTIHEFKSAKTKLLIKIDPALKIKTKKVTADOCANRCTRNKGLPFTCAVFDK 60
Db 1 QRRKNTIHEFKSAKTKLLIKIDPALKIKTKKVTADOCANRCTRNKGLPFTCAVFDK 60
Qy 61 ARKOCCLMPFNSMSGYKKEFGHEFDLYENKDYIRNCLIGKNSYKGTVSITKSGIKQIP 120
Db 61 ARKOCCLMPFNSMSGYKKEFGHEFDLYENKDYIRNCLIGKNSYKGTVSITKSGIKQIP 120
Qy 121 WSSMIPEHSHFLPSYRGKDLQENYCRNPGEGSPWCTSNPEVRVECDIPCCSEYEC 180
Db 121 WSSMIPEHSHFLPSYRGKDLQENYCRNPGEGSPWCTSNPEVRVECDIPCCSEYEC 180
Qy 181 MTCNGESYRGIMDHTESGKICQRMHDQTPHHRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240
Db 181 MTCNGESYRGIMDHTESGKICQRMHDQTPHHRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240
Qy 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGVTNIMGIFCQWRDS 300
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGVTNIMGIFCQWRDS 300
Qy 301 QYPHEHMTPEPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 301 QYPHEHMTPEPFCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
Qy 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420
Db 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420
Qy 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

RESULT 6

AAU04275
ID AAU04275 standard; peptide; 697 AA.

XX AAU04275;

XX 23-OCT-2001 (first entry)

XX Nuclear ligand Peg20 used in nucleic acid transporter system.

XX Nucleic acid transport; cytosols; ligand; lysis agent; spacer molecule;
KW gene therapy; hepatocyte; muscle; bone forming cell.

XX Synthetic.

XX US6177554-B1.

XX 23-JAN-2001.

XX 05-JUN-1995; 95US-00462040.

XX 20-MAR-1992; 92US-00855389.

XX 19-MAR-1993; 93WO-US002725.

XX 14-DEC-1993; 93US-00167641.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Woo SLC, Smith LC, Cristiano RJ, Gottchalk S, Sparrow J;

XX WPI; 2001-365933/38.

XX Nucleic acid transporter system, useful for creating transgenic animals for
PT assessing human disease such as cancer in an animal model.

XX XX Disclosure; Fig 23A; 11pp; English.
PS XX
CC The sequence represents the nuclear ligand, Pep20, used in a nucleic acid
CC transporter system. The nucleic acid transporter system uses nucleic acid
CC binding complexes containing surface ligands which are capable of binding
CC to a cell surface receptor and entering the cell through cytolysis. The
CC compounds of the invention are either ligands, binding molecules (surface
CC ligands), lysis agents, spacer molecules or their intermediates. The
CC ligands, binding molecules, lysis agents and spacer molecules are used in
CC nucleic acid transporter systems to deliver nucleic acid into specific
CC cells e.g. in gene therapy to deliver nucleic acid into hepatocytes,
CC muscle cells or bone forming cells
XX
SQ Sequence 697 AA;
Query Match 99.9%; Score 2601; DB 4; Length 697;
Best Local Similarity 99.8%; Pred. No. 8.7e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNVTADQCANRCTRNGKLPFTCKAFVFDK 60
DB 1 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNVTADQCANRCTRNGKLPFTCKAFVFDK 60
QY 61 ARKQCIWPFPPNSMSGVKKFEGHEFDLYENKDIYINNCIIIGKGRSYKGTVSITKSGIKCOP 120
DB 61 ARKQCIWPFPPNSMSGVKKFEGHEFDLYENKDIYINNCIIIGKGRSYKGTVSITKSGIKCOP 120
QY 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNREGGEGWCTSNPEVRYEVCDI PQCESEVFC 180
DB 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNREGGEGWCTSNPEVRYEVCDI PQCESEVFC 180
QY 181 MTCNGESYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNRYCRNPDGQPRPMC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNRYCRNPDGQPRPMC 240
QY 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGVTNTIWNIGIPCORWDS 300
DB 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGVTNTIWNIGIPCORWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGDGC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGDGC 360
QY 361 YGNGKNYVGNLSQTRSGLTCSMDKMMEDLHRHIFWEPDASKLBNYCRNPDGDAHGM 420
DB 361 YGNGKNYVGNLSQTRSGLTCSMDKMMEDLHRHIFWEPDASKLBNYCRNPDGDAHGM 420
QY 421 CYTGNPLIPWDYCPISRCEGDTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCEGDTPTIV 447
RESULT 7
ID AAR25676
AA AAR25676 standard; protein; 728 AA.
AC AAR25676;
XX
XX 20-JAN-1993 (first entry)
DE Recombinant human hepatocyte growth factor.
XX
XX HGF; enhance growth; preparing transgenic animals; hepatic disease;
KW clinical diagnostic reagent; drug.
XX
XX Homo sapiens.
OS
XX
XX JP04183394-A.
XX PN
XX 30-JUN-1992.
XX PD
XX 19-NOV-1990; 90JP-00314548.
XX PF

XX XX 19-NOV-1990; 90JP-00314548.
PR XX
XX (TOYMA) TOYOBO KK.
PA (NAKA/) NAKAMURA T.
XX
XX WPI; 1992-265591/32.
DR
XX
XX Recombinant human hepatocyte growth factor and DNA encoding it - useful
PT for diagnosis and treatment of hepatic disease and transgenic animal
PT prepn.
XX
XX Disclosure; Page 11; 28pp; Japanese.
PS
XX
XX This sequence represents a recombinant human hepatocyte growth factor. It
CC has physiological activity, and using it enhanced growth of hepatocytes
CC is possible. It is useful as a clinical diagnostic reagent, or a drug for
CC treating hepatic disease. See also AAR25676-92, AAO26713-27
CC
XX
SQ Sequence 728 AA;
Query Match 99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNVTADQCANRCTRNGKLPFTCKAFVFDK 60
DB 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNVTADQCANRCTRNGKLPFTCKAFVFDK 91
QY 61 ARKQCIWPFPPNSMSGVKKFEGHEFDLYENKDIYINNCIIIGKGRSYKGTVSITKSGIKCOP 120
DB 61 ARKQCIWPFPPNSMSGVKKFEGHEFDLYENKDIYINNCIIIGKGRSYKGTVSITKSGIKCOP 151
QY 92 ARKQCIWPFPPNSMSGVKKFEGHEFDLYENKDIYINNCIIIGKGRSYKGTVSITKSGIKCOP 151
DB 92 ARKQCIWPFPPNSMSGVKKFEGHEFDLYENKDIYINNCIIIGKGRSYKGTVSITKSGIKCOP 151
QY 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNREGGEGWCTSNPEVRYEVCDI PQCESEVFC 180
DB 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNREGGEGWCTSNPEVRYEVCDI PQCESEVFC 211
QY 152 WSSMI PHEHSFLPSSYRGKDLQENYCRNREGGEGWCTSNPEVRYEVCDI PQCESEVFC 211
DB 152 WSSMI PHEHSFLPSSYRGKDLQENYCRNREGGEGWCTSNPEVRYEVCDI PQCESEVFC 211
QY 181 MTCNGESYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNRYCRNPDGQPRPMC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNRYCRNPDGQPRPMC 271
QY 212 MTCNGESYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNRYCRNPDGQPRPMC 271
DB 212 MTCNGESYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNRYCRNPDGQPRPMC 271
QY 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGVTNTIWNIGIPCORWDS 300
DB 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGVTNTIWNIGIPCORWDS 331
QY 272 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGVTNTIWNIGIPCORWDS 331
DB 272 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGVTNTIWNIGIPCORWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGDGC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGDGC 391
QY 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGDGC 391
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGDGC 391
QY 361 YGNGKNYVGNLSQTRSGLTCSMDKMMEDLHRHIFWEPDASKLBNYCRNPDGDAHGM 420
DB 361 YGNGKNYVGNLSQTRSGLTCSMDKMMEDLHRHIFWEPDASKLBNYCRNPDGDAHGM 451
QY 392 YGNGKNYVGNLSQTRSGLTCSMDKMMEDLHRHIFWEPDASKLBNYCRNPDGDAHGM 451
DB 392 YGNGKNYVGNLSQTRSGLTCSMDKMMEDLHRHIFWEPDASKLBNYCRNPDGDAHGM 451
RESULT 8
ID AAR20005
AA AAR20005 standard; protein; 728 AA.
AC AAR20005;
XX
XX 24-MAR-1992 (first entry)
DE Human hepatocyte growth factor.
XX
XX HGF; kidney regeneration; nephritis.
KW
XX
XX Homo sapiens.
OS
XX
XX EP462549-A.
XX PN

```
XX 27-DEC-1991.
PD 18-JUN-1991; 91EP-00109923.
XX 18-JUN-1991; 91EP-00109923.
XX 19-JUN-1990; 90JP-00158841.
XX (TOYM ) TOYO BOSEKI KK.
XX Nakamura T;
XX WPI; 1992-000939/01.
DR N-PSDB; AAQ20049.
XX Agent contg. hepatocyte growth factor and carrier - used for treating
PT renal diseases and promoting nephrocyte growth and as diagnostic for
PT renal diseases.
XX Disclosure; Fig 1; 15pp; English.
XX Human HGF comprises an alpha-chain of 440 amino acids and a beta-chain of
CC 234 amino acids. There are 4 kringe domains in the alpha-chain, similar
CC to that of plasmin; the beta-chain has about 37 per cent homology with
CC the beta-chain of plasmin having serine protease activity. Homology of
CC the amino acid sequence of rat HGF and human HGF is 91.6 per cent in the
CC alpha-chain and 88.9 per cent in the beta-chain. HGF has been found to be
CC an agent for nephrocyte growth and is useful as a treatment for renal
CC failure
XX
XX Sequence 728 AA;
SQ
Query Match 99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRRNTTHERFKSAKTLIKIDPALKITKKVNTADOCANRCTRNKGLPTCKAFVFDK 60
DB :|||||
DB 32 QKRKRNTTHERFKSAKTLIKIDPALKITKKVNTADOCANRCTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKOP 120
DB :|||||
DB 92 ARKQCLMPFPNMSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKOP 151
QY 121 WSSMTIPHESFLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVREVCIDIPQSEVEVC 180
DB :|||||
DB 152 WSSMTIPHESFLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVREVCIDIPQSEVEVC 211
QY 181 MTCNGBSYRGLMDHTESGKICQRMHQTPHRHKLPERYPDKGFDNFCRNPDGQPRPMC 240
DB :|||||
DB 212 MTCNGBSYRGLMDHTESGKICQRMHQTPHRHKLPERYPDKGFDNFCRNPDGQPRPMC 271
QY 241 YTLDPHTRWYCAITKCANNTMNDVPLETTECIQGGGEGYRGVNTIWNIGIPQCRWMS 300
DB :|||||
DB 272 YTLDPHTRWYCAITKCANNTMNDVPLETTECIQGGGEGYRGVNTIWNIGIPQCRWMS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
DB :|||||
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRNGNKNTYNGNISQTRSGITCSWMDKNMEDLHRHIFWEPDASKLLENYCRNPDGDAHGW 420
DB :|||||
DB 392 YRNGNKNTYNGNISQTRSGITCSWMDKNMEDLHRHIFWEPDASKLLENYCRNPDGDAHGW 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTPTTV 447
DB :|||||
DB 452 CYTGNPLIPWDYCPISRCGEGDTPTTV 478
RESULT 9
AAR40862
ID AAR40862 standard; protein; 728 AA.
XX
AC AAR40862;
```

```
XX 14-MAR-1994 (first entry)
DT Competitive inhibitor of HGF.
XX
DE HGF; hepatocyte growth factor; transformation; antagonistic; liver disease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..31
XX Protein /label= sig_peptide
XX /label= 32..728
XX /label= mat_protein
XX Misc-difference 494
XX /note= "site of possible mutation"
XX
XX JP05208998-A.
XX
XX 20-AUG-1993.
XX
XX 25-DEC-1991; 91JP-00357040.
XX
XX 25-DEC-1991; 91JP-00357040.
XX
XX (MITU ) MITSUBISHI KASEI CORP.
XX
XX WPI; 1993-297806/38.
XX N-PSDB; AAQ47832.
XX
XX New protein is competitive inhibitor of hepatocyte growth factor - used
XX in animal model of liver disease.
XX
XX Claim 7; Page 8-11; 20pp; Japanese.
XX
XX The gene encodes an inhibitor of HGF, which has a molecular weight of ca.
XX 76000 to 92000 as determined by SDS-PAGE. Human modified (NP) HGF-A and
XX B fragments were prepared from human HGF cDNA. This was amplified by PCR
XX to give a human NP-HGF DNA fragment that was transformed into E.coli.
XX (see AAQ47833) for a related sequence
XX
XX Sequence 728 AA;
SQ
Query Match 99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRRNTTHERFKSAKTLIKIDPALKITKKVNTADOCANRCTRNKGLPTCKAFVFDK 60
DB :|||||
DB 32 QKRKRNTTHERFKSAKTLIKIDPALKITKKVNTADOCANRCTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKOP 120
DB :|||||
DB 92 ARKQCLMPFPNMSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKOP 151
QY 121 WSSMTIPHESFLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVREVCIDIPQSEVEVC 180
DB :|||||
DB 152 WSSMTIPHESFLPSSYRGKDLQENYCRNPRGEGGPMCTSNPEVREVCIDIPQSEVEVC 211
QY 181 MTCNGBSYRGLMDHTESGKICQRMHQTPHRHKLPERYPDKGFDNFCRNPDGQPRPMC 240
DB :|||||
DB 212 MTCNGBSYRGLMDHTESGKICQRMHQTPHRHKLPERYPDKGFDNFCRNPDGQPRPMC 271
QY 241 YTLDPHTRWYCAITKCANNTMNDVPLETTECIQGGGEGYRGVNTIWNIGIPQCRWMS 300
DB :|||||
DB 272 YTLDPHTRWYCAITKCANNTMNDVPLETTECIQGGGEGYRGVNTIWNIGIPQCRWMS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
DB :|||||
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRNGNKNTYNGNISQTRSGITCSWMDKNMEDLHRHIFWEPDASKLLENYCRNPDGDAHGW 420
DB :|||||
```

```
Db      392 YRGNKKNYGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDDAHGPM 451
QY      421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447
      |||
Db      452 CYTGNPLIPWDYCPISRCEGDTTPTIV 478

RESULT 10
AAR40863
ID      AAR40863 standard; protein; 728 AA.
XX
AC      AAR40863;
XX
DT      14-MAR-1994 (first entry)
XX
DE      Competitive inhibitor of HGF.
XX
KM      HGF, hepatocyte growth factor; transformation; antagonist; liver disease.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      Peptide 1..31
FT      Protein /label= sig_peptide
FT      /label= mat_protein
FT      Misc-difference 489
FT      /note= "site of possible mutation"
FT      Misc-difference 491
FT      /note= "site of possible mutation"
FT      Misc-difference 494
FT      /note= "site of possible mutation"
XX
PN      JP05208998-A.
XX
PD      20-AUG-1993.
XX
PF      25-DEC-1991; 91JP-00357040.
XX
PR      25-DEC-1991; 91JP-00357040.
XX
PA      (MITU ) MITSUBISHI KASEI CORP.
XX
DR      WPI; 1993-297806/38.
XX
DR      N-PSDB; AAQ47833.
XX
PT      New protein is competitive inhibitor of hepatocyte growth factor - used
PT      in animal model of liver disease.
XX
PS      Claim 6; Page 11-14; 20pp; Japanese.
XX
CC      The gene encodes an inhibitor of HGF, which has a molecular weight of ca.
CC      76000 to 92000 as determined by SDS-PAGE. Human modified (NP) HGF-A and -
CC      B fragments were prepared from human HGF CDNA. This was amplified by PCR
CC      to give a human NP-HGF DNA fragment that was transformed into E.coli.
CC      (see AAQ47833) for a related sequence
XX
SQ      Sequence 728 AA:

Query Match      99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 EKKRRNTHEPKRSATTTLIKIDPAKIKTKKYNVTADOCANRCTRNGKLPFTCKAFVFK 60
      |||
Db      32 QKRRNTHEPKRSATTTLIKIDPAKIKTKKYNVTADOCANRCTRNGKLPFTCKAFVFK 91
      |||
QY      61 ARKQCLMPFPNMSGVKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKOP 120
      |||
Db      92 ARKQCLMPFPNMSGVKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKOP 151
      |||
QY      121 WSGMIFHEHSFLPSSYRGKDLQENYCRNPGEEGGFWCTSNPEVREYECIDIPOCSEVEC 180
      |||
```

```
Db      152 WSGMIFHEHSFLPSSYRGKDLQENYCRNPGEEGGFWCTSNPEVREYECIDIPOCSEVEC 211
QY      181 MTCNGESYRGKDLMDHTESGKICORMDHQTPHRKFLPERYPDGFDNNYCRNPDGQRPWC 240
      |||
Db      212 MTCNGESYRGKDLMDHTESGKICORMDHQTPHRKFLPERYPDGFDNNYCRNPDGQRPWC 271
      |||
QY      241 YTLDPHTMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNIIWNGIPCORMDS 300
      |||
Db      272 YTLDPHTMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNIIWNGIPCORMDS 331
      |||
QY      301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESWCTTDPNIRVYCSQIPNCDSHGQDC 360
      |||
Db      332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESWCTTDPNIRVYCSQIPNCDSHGQDC 391
      |||
QY      361 YRGNKKNYGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDDAHGPM 420
      |||
Db      392 YRGNKKNYGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDDAHGPM 451
      |||
QY      421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447
      |||
Db      452 CYTGNPLIPWDYCPISRCEGDTTPTIV 478

RESULT 11
AAR42062
ID      AAR42062 standard; protein; 728 AA.
XX
AC      AAR42062;
XX
DT      25-MAR-2003 (revised)
DT      15-NOV-1993 (first entry)
XX
DE      Vascular endothelial cell growth enhancer protein.
XX
KM      Enhance; growth; vascular endothelial cell; human; tumour; cell line;
KM      HUOCA-II; HUOCA-III; blood vessel; wounds; burns; decubitus;
KM      post-operative tissue damage; drug; cardiac angiopathy.
XX
OS      Homo sapiens.
XX
PN      EP550296-A2.
XX
PD      07-JUL-1993.
XX
PF      27-NOV-1992; 92EP-00403199.
XX
PR      28-NOV-1991; 91JP-00337999.
XX
PA      (TERU ) TERUMO CORP.
XX
PI      Sudo T, Harada K, Hirahara I, Adachi M;
XX
DR      WPI; 1993-215669/27.
XX
DR      N-PSDB; AAQ45702.
XX
PT      Vascular endothelial cell growth factor protein - used for promoting
PT      angiogenesis in the treatment of cardiac angiopathy, wounds, burn
PT      injuries, postoperative tissue damage etc.
XX
PS      Claim 6; Page 18-21; 44pp; English.
XX
CC      This sequence represents a single chain protein which selectively
CC      enhances the growth of vascular endothelial cells. This protein was
CC      produced by the human tumour cell line HUOCA-II or HUOCA-III. This
CC      protein enhances the formation of new blood vessels and may be used to
CC      enhance healing of wounds, burns injuries, decubitus or post-operative
CC      tissue damage. It may also be used as a drug for cardiac angiopathy.
CC      (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ      Sequence 728 AA:

Query Match      99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
```

```
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRNTIHEPKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVFDK 60
DB 32 QKRKNTIHEPKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPMSSGCVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCP 120
DB 92 ARKQCLMPFPMSSGCVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCP 151
QY 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDI PQCEVEEC 180
DB 152 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDI PQCEVEEC 211
QY 181 MTCNGESYRGMLMDHTESGKICQRMWDQTPHRAKFLPERYPDGFPDNYCRNPDGQPRPWC 240
DB 212 MTCNGESYRGMLMDHTESGKICQRMWDQTPHRAKFLPERYPDGFPDNYCRNPDGQPRPWC 271
QY 241 YTLDPHTWEXCAITKCADNTMNDVPLETTECIIQGSGEYRGTVNTIMNGIPCORWDS 300
DB 272 YTLDPHTWEXCAITKCADNTMNDVPLETTECIIQGSGEYRGTVNTIMNGIPCORWDS 331
QY 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 332 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRGNGKNYMGNLSTQRSGLTCSMDKXMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 420
DB 392 YRGNGKNYMGNLSTQRSGLTCSMDKXMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCEGDTTPTIV 478

RESULT 12
AAR87522
ID AAR87522 standard; protein: 728 AA.
AC AAR87522;
XX
DT 21-JUN-1996 (first entry)
DE Mutant hepatocyte growth factor #1 with changes at residues 491-495.
KW Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa;
KM injury; blood vessel; point mutation.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 491..495
FT /note= "change from wild type sequence: Lys-Glu-Leu-Arg-
FT Val to Ile-Glu-Gly-Arg-Thr"
XX
XX JP07304796-A.
XX
XX 21-NOV-1995.
XX
XX 07-MAY-1994; 94JP-00117506.
XX
XX 07-MAY-1994; 94JP-00117506.
XX
XX (TERU ) TERUMO CORP.
XX
XX WPI; 1996-035890/04.
XX
XX Novel derivative of hepatocyte growth factor - comprises a mutation at
XX positions 492-494, for use in treating injured blood vessels.
XX
XX Disclosure; Page; 8pp; Japanese.
XX
XX The amino acid of the hepatocyte growth factor (HGF) generated to contain
```

```
CC amino acid changes at residues 492-494. (Note - this sequence is not
CC given in the specification but is based on the HGF sequence given in
CC AAR0005). The generated mutant contains the sequence Ile-Glu-Gly-Arg-
CC Thr, which replaces the wild type sequence Lys-Glu-Leu-Arg-Val. The HGF
CC is translated as single chain protein which is activated to a two chain
CC protein by protease cleavage between residues 494-5, producing the alpha
CC and beta chains. The novel sequence alters the protease recognition site,
CC putatively introducing a Factor Xa recognition site. The novel HGF deriv.
CC can be activated specifically at the site of injury in a blood vessel.
CC (See AAR87522-5 for different HGF derivs.). The novel HGF derivs. were
CC generated by point mutations using the oligonucleotides AAT06762-5
XX
SQ Sequence 728 AA;
XX
Query Match 99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRNTIHEPKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVFDK 60
DB 32 QKRKNTIHEPKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPMSSGCVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCP 120
DB 92 ARKQCLMPFPMSSGCVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCP 151
QY 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDI PQCEVEEC 180
DB 152 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDI PQCEVEEC 211
QY 181 MTCNGESYRGMLMDHTESGKICQRMWDQTPHRAKFLPERYPDGFPDNYCRNPDGQPRPWC 240
DB 212 MTCNGESYRGMLMDHTESGKICQRMWDQTPHRAKFLPERYPDGFPDNYCRNPDGQPRPWC 271
QY 241 YTLDPHTWEXCAITKCADNTMNDVPLETTECIIQGSGEYRGTVNTIMNGIPCORWDS 300
DB 272 YTLDPHTWEXCAITKCADNTMNDVPLETTECIIQGSGEYRGTVNTIMNGIPCORWDS 331
QY 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 332 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRGNGKNYMGNLSTQRSGLTCSMDKXMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 420
DB 392 YRGNGKNYMGNLSTQRSGLTCSMDKXMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCEGDTTPTIV 478

RESULT 13
AAR87524
ID AAR87524 standard; protein: 728 AA.
AC AAR87524;
XX
DT 21-JUN-1996 (first entry)
DE Mutant hepatocyte growth factor #3 with changes at residues 491-495.
KW Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa;
KM injury; blood vessel; point mutation.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 491..495
FT /note= "change from wild type sequence: Lys-Glu-Leu-Arg-
FT Val to Lys-Glu-Gly-Arg-Thr"
XX
XX JP07304796-A.
```

PD 21-NOV-1995.
XX
XX 07-MAY-1994; 94JP-00117506.
XX
PR 07-MAY-1994; 94JP-00117506.
XX
PA (TERU) TERUMO CORP.
XX
XX WPI; 1996-035890/04.
XX
PT Novel derivative of hepatocyte growth factor - comprises a mutation at
XX positions 492-494, for use in treating injured blood vessels.
XX
XX Disclosure; Page; 8pp; Japanese.
XX
CC The amino acid of the hepatocyte growth factor (HGF) generated to contain
CC amino acid changes at residues 492-494. (Note - this sequence is not
CC given in the specification but is based on the HGF sequence given in
CC AAR20005). The generated mutant contains the sequence Lys-Gln-Gly-Arg-
CC Thr, which replaces the wild type sequence Lys-Glu-Leu-Arg-Val. The HGF
CC is translated as single chain protein which is activated to a two chain
CC protein by protease cleavage between residues 494-5, producing the alpha
CC and beta chains. The novel sequence alters the protease recognition site,
CC putatively introducing a Factor Xa recognition site. The novel HGF deriv.
CC (See AAR87522-5 for different HGF derivs.). The novel HGF derivs. were
CC generated by point mutations using the oligonucleotides AAT06762-5
XX
XX Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERRRRNTIIEFKSAKTTLIKIDPAKIKTKKVNVTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 QRRRRNTIIEFKSAKTTLIKIDPAKIKTKKVNVTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 120
DB 92 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 151
QY 121 WSSMIPHEHSFLPSSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCEVEVC 180
DB 152 WSSMIPHEHSFLPSSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCEVEVC 211
QY 181 MTCNGSSYRGKDLMDHTESGKICQRMWDQTPHRRKFLPERYPDGFDPNYCRNPDGQRPWC 240
DB 212 MTCNGSSYRGKDLMDHTESGKICQRMWDQTPHRRKFLPERYPDGFDPNYCRNPDGQRPWC 271
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIQOGGEGYRGVTNTIWNIGIPCGRWD 300
DB 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIQOGGEGYRGVTNTIWNIGIPCGRWD 331
QY 301 QYPEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 332 QYPEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRNGNGKNTMGNLSQTSGLTCSWMDXKMDLHRHITWEPPASXLNENYCRNPDGDAHGPW 420
DB 392 YRNGNGKNTMGNLSQTSGLTCSWMDXKMDLHRHITWEPPASXLNENYCRNPDGDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCBGDTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCBGDTPTIV 478

RESULT 14
AAR87525
ID AAR87525 standard; protein; 728 AA.
XX
XX AAR87525;
XX

DT 21-JUN-1996 (first entry)
XX
XX Mutant hepatocyte growth factor #4 with changes at residues 491-495.
DE
XX Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa;
XX injury; blood vessel; point mutation.
KM
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 491..495
FT /note="change from wild type sequence: Lys-Glu-Leu-Arg-
FT val to Lys-Gln-Gly-Arg-Val"
XX
XX JP07304796-A.
XX
PD 21-NOV-1995.
XX
XX 07-MAY-1994; 94JP-00117506.
XX
XX 07-MAY-1994; 94JP-00117506.
XX
XX (TERU) TERUMO CORP.
XX
XX WPI; 1996-035890/04.
XX
XX Novel derivative of hepatocyte growth factor - comprises a mutation at
XX positions 492-494, for use in treating injured blood vessels.
XX
XX Disclosure; Page; 8pp; Japanese.
XX
XX The amino acid of the hepatocyte growth factor (HGF) generated to contain
XX amino acid changes at residues 492-494. (Note - this sequence is not
XX given in the specification but is based on the HGF sequence given in
XX AAR20005). The generated mutant contains the sequence Lys-Gln-Gly-Arg-
XX Val, which replaces the wild type sequence Lys-Glu-Leu-Arg-Val. The HGF
XX is translated as single chain protein which is activated to a two chain
XX protein by protease cleavage between residues 494-5, producing the alpha
XX and beta chains. The novel sequence alters the protease recognition site,
XX putatively introducing a Factor Xa recognition site. The novel HGF deriv.
XX (See AAR87522-5 for different HGF derivs.). The novel HGF derivs. were
XX generated by point mutations using the oligonucleotides AAT06762-5
XX
XX Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 2; Length 728;
Best Local Similarity 99.8%; Pred. No. 9.1e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERRRRNTIIEFKSAKTTLIKIDPAKIKTKKVNVTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 QRRRRNTIIEFKSAKTTLIKIDPAKIKTKKVNVTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 120
DB 92 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 151
QY 121 WSSMIPHEHSFLPSSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCEVEVC 180
DB 152 WSSMIPHEHSFLPSSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCDDIPQCEVEVC 211
QY 181 MTCNGSSYRGKDLMDHTESGKICQRMWDQTPHRRKFLPERYPDGFDPNYCRNPDGQRPWC 240
DB 212 MTCNGSSYRGKDLMDHTESGKICQRMWDQTPHRRKFLPERYPDGFDPNYCRNPDGQRPWC 271
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIQOGGEGYRGVTNTIWNIGIPCGRWD 300
DB 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIQOGGEGYRGVTNTIWNIGIPCGRWD 331
QY 301 QYPEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 332 QYPEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDSHGQDC 391

QY 361 YRNGKNYVGNLSQTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNDDDAHGPM 420
 DB 392 YRNGKNYVGNLSQTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNDDDAHGPM 451
 QY 421 CYTGNPLIPMDYCPISRCGDTTPTIV 447
 DB 452 CYTGNPLIPMDYCPISRCGDTTPTIV 478

RESULT 15

AAR87523 standard; protein; 728 AA.

AC AAR87523;

DT 21-JUN-1996 (first entry)

DE Mutant hepatocyte growth factor #2 with changes at residues 491-495.

KW Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa; injury; blood vessel; point mutation.

OS Synthetic.

FX Key Location/Qualifiers

FT Misc-difference 491..495

FT /note="change from wild type sequence: Lys-Glu-Leu-Arg-Val to Lys-Glu-Gly-Arg-Ile"

PN JP07304796-A.

PD 21-NOV-1995.

PF 07-MAY-1994; 94JP-00117506.

PR 07-MAY-1994; 94JP-00117506.

PA (TERU) TERUMO CORP.

WPI; 1996-035890/04.

PT Novel derivative of hepatocyte growth factor - comprises a mutation at positions 492-494, for use in treating injured blood vessels.

PS Disclosure; Page; 8pp; Japanese.

CC The amino acid of the hepatocyte growth factor (HGF) generated to contain amino acid changes at residues 492-494. (Note - this sequence is not given in the specification but is based on the HGF sequence given in CC AAR20005). The generated mutant contains the sequence Lys-Glu-Gly-Arg-Ile, which replaces the wild type sequence Lys-Glu-Leu-Arg-Val. The HGF is translated as single chain protein which is activated to a two chain protein by protease cleavage between residues 494-5, producing the alpha and beta chains. The novel sequence alters the protease recognition site, putatively introducing a Factor Xa recognition site. The novel HGF deriv. can be activated specifically at the site of injury in a blood vessel. (See AAR87522-5 for different HGF derivs.). The novel HGF derivs. were generated by point mutations using the oligonucleotides AAR06762-5

SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 2; Length 728;

Best Local Similarity 99.8%; Pred. No. 9..1e-165; Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIEFKKSATTTIKIDPAKIKTKKVVNTADOCANRCTRNKGLPFTCKAFVFDK 60
 DB 32 QRRRRNTIEFKKSATTTIKIDPAKIKTKKVVNTADOCANRCTRNKGLPFTCKAFVFDK 91
 QY 61 ARKOCILPFPNMSGVKKKPGHEPDLLENKDYIRNCIIIGKSGYKGVSIITKSGIKCOP 120
 DB 92 ARKOCILPFPNMSGVKKKPGHEPDLLENKDYIRNCIIIGKSGYKGVSIITKSGIKCOP 151

QY 121 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFISNPEVREVCDIPOCSEVEC 180
 DB 152 WSSMI PHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFISNPEVREVCDIPOCSEVEC 211
 QY 181 MTCNGESYRGIMDHTESGKICQWRDHOPIRHHKFLPERYPDKGFPDNYCRNPDQPRPMC 240
 DB 212 MTCNGESYRGIMDHTESGKICQWRDHOPIRHHKFLPERYPDKGFPDNYCRNPDQPRPMC 271
 QY 241 YTLDPHTWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGTVNTINMGIPQQRWDS 300
 DB 272 YTLDPHTWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGTVNTINMGIPQQRWDS 331
 QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDMSHGDC 360
 DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDMSHGDC 391
 QY 361 YRNGKNYVGNLSQTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNDDDAHGPM 420
 DB 392 YRNGKNYVGNLSQTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNDDDAHGPM 451
 QY 421 CYTGNPLIPMDYCPISRCGDTTPTIV 447
 DB 452 CYTGNPLIPMDYCPISRCGDTTPTIV 478

Search completed: October 3, 2005, 06:05:50
 Job time : 174.465 secs


```
Db 61 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDIYINCLIIIGKRSYKGVSTITKSGIKOP 120
Qy 121 WSSMT.PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEYREYECDIPOCSEVEC 180
Db 121 WSSMT.PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEYREYECDIPOCSEVEC 180
Qy 181 MTCNGESTYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFPDNYCRNPDGQPRPWC 240
Db 181 MTCNGESTYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFPDNYCRNPDGQPRPWC 240
Qy 241 YTLDPHTHREYCAIKTCADNTMNDTVPLETTTECLOGGEGYRGTVNTIWMGIPCORWDS 300
Db 241 YTLDPHTHREYCAIKTCADNTMNDTVPLETTTECLOGGEGYRGTVNTIWMGIPCORWDS 300
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
Qy 361 YRGNGKNYMGUNLSQTRSGLTCSMMDKXNEDLHRHIFWEPDASKLBNENYCRNPDGDAHGFW 420
Db 361 YRGNGKNYMGUNLSQTRSGLTCSMMDKXNEDLHRHIFWEPDASKLBNENYCRNPDGDAHGFW 420
Qy 421 CYTGNPLIPWDYCPISRCGSDTTPITV 447
Db 421 CYTGNPLIPWDYCPISRCGSDTTPITV 447
```

```
RESULT 2
US-10-081-309-2
; Sequence 2, Application US/10081309
; Publication No. US20030012775A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Peg Conjugates of NK4
; FILE REFERENCE: 20859
; CURRENT APPLICATION NUMBER: US/10/081,309
; CURRENT FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 447
; TYPE: PRT
; ORGANISM: homosapiens
US-10-081-309-2
```

```
Query Match 99.9%; Score 2601; DB 14; Length 447;
Best Local Similarity 99.8%; Pred. No. 1.1e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 EKKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPTCKAFVFDK 60
Db 1 EKKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPTCKAFVFDK 60
Qy 61 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDIYINCLIIIGKRSYKGVSTITKSGIKOP 120
Db 61 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDIYINCLIIIGKRSYKGVSTITKSGIKOP 120
Qy 121 WSSMT.PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEYREYECDIPOCSEVEC 180
Db 121 WSSMT.PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEYREYECDIPOCSEVEC 180
Qy 181 MTCNGESTYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFPDNYCRNPDGQPRPWC 240
Db 181 MTCNGESTYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFPDNYCRNPDGQPRPWC 240
Qy 241 YTLDPHTHREYCAIKTCADNTMNDTVPLETTTECLOGGEGYRGTVNTIWMGIPCORWDS 300
Db 241 YTLDPHTHREYCAIKTCADNTMNDTVPLETTTECLOGGEGYRGTVNTIWMGIPCORWDS 300
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
```

```
Qy 361 YRGNGKNYMGUNLSQTRSGLTCSMMDKXNEDLHRHIFWEPDASKLBNENYCRNPDGDAHGFW 420
Db 361 YRGNGKNYMGUNLSQTRSGLTCSMMDKXNEDLHRHIFWEPDASKLBNENYCRNPDGDAHGFW 420
Qy 421 CYTGNPLIPWDYCPISRCGSDTTPITV 447
Db 421 CYTGNPLIPWDYCPISRCGSDTTPITV 447
```

```
RESULT 3
US-10-926-088A-1
; Sequence 1, Application US/10926088A
; Publication No. US20050164918A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA Toshikazu
; APPLICANT: MATSUMOTO Kunio
; APPLICANT: FUKUTA Kazuhito
; TITLE OF INVENTION: A segment of glycosylation-deficient HGF alpha-chain
; FILE REFERENCE: 2004-13284/W/C/01736
; CURRENT APPLICATION NUMBER: US/10/926,088A
; CURRENT FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 1
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: alpha-chain of human hepatocyte growth factor
US-10-926-088A-1
```

```
Query Match 99.9%; Score 2601; DB 18; Length 494;
Best Local Similarity 99.8%; Pred. No. 1.3e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 EKKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPTCKAFVFDK 60
Db 32 EKKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPTCKAFVFDK 91
Qy 61 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDIYINCLIIIGKRSYKGVSTITKSGIKOP 120
Db 92 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDIYINCLIIIGKRSYKGVSTITKSGIKOP 151
Qy 121 WSSMT.PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEYREYECDIPOCSEVEC 180
Db 152 WSSMT.PHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEYREYECDIPOCSEVEC 211
Qy 181 MTCNGESTYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFPDNYCRNPDGQPRPWC 240
Db 212 MTCNGESTYRGKDLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFPDNYCRNPDGQPRPWC 271
Qy 241 YTLDPHTHREYCAIKTCADNTMNDTVPLETTTECLOGGEGYRGTVNTIWMGIPCORWDS 300
Db 272 YTLDPHTHREYCAIKTCADNTMNDTVPLETTTECLOGGEGYRGTVNTIWMGIPCORWDS 331
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 391
Qy 361 YRGNGKNYMGUNLSQTRSGLTCSMMDKXNEDLHRHIFWEPDASKLBNENYCRNPDGDAHGFW 420
Db 392 YRGNGKNYMGUNLSQTRSGLTCSMMDKXNEDLHRHIFWEPDASKLBNENYCRNPDGDAHGFW 451
Qy 421 CYTGNPLIPWDYCPISRCGSDTTPITV 447
Db 452 CYTGNPLIPWDYCPISRCGSDTTPITV 478
```

```
RESULT 4
US-10-872-198-120
; Sequence 120, Application US/10872198
; Publication No. US20050002897A1
; GENERAL INFORMATION:
; APPLICANT: Ulrich HAUPTS
```



```
; APPLICANT: Andre KOLTERMANN
; APPLICANT: Andreas SCHEIDIG
; APPLICANT: Christian VOTSMERER
; APPLICANT: Ulrich Ketting
; TITLE OF INVENTION: NEW BIOLOGICAL ENTITIES AND USE THEREOF
; FILE REFERENCE: 04156.000204
; CURRENT APPLICATION NUMBER: US/10/872.198
; CURRENT FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 60/543,518
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/524,960
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: EP 04003058
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: EP 03025871
; PRIOR FILING DATE: 2003-11-11
; PRIOR APPLICATION NUMBER: EP 03025851
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: EP 03013819
; PRIOR FILING DATE: 2003-06-18
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-872-198-120
```

```
Query Match          99.9%; Score 2601; DB 17; Length 726;
Best Local Similarity 99.8%; Pred. No. 2e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ERRKRRYTHIEFKKSATTLIKIDPALKIKTKKVNADQCANRCTRNGKLPFTCKAFVFDK 60
      :
DB      30 QRRRRNTIHEFKKSATTLIKIDPALKIKTKKVNADQCANRCTRNGKLPFTCKAFVFDK 89
QY      61 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDYIRNCIIGKGSYKGTVAITKSGIKCOP 120
      :
DB      90 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDYIRNCIIGKGSYKGTVAITKSGIKCOP 149
QY      121 WSSMIDHESFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNDEVRVEVCDIPQCESEVEC 180
      :
DB      150 WSSMIDHESFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNDEVRVEVCDIPQCESEVEC 209
QY      181 MTCNGSSYRGKDLMDHTSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRRPWC 240
      :
DB      210 MTCNGSSYRGKDLMDHTSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRRPWC 269
QY      241 YTLDPHTRMWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNGLPCQRMDS 300
      :
DB      270 YTLDPHTRMWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNGLPCQRMDS 329
QY      301 QYHEHDMTPENKCKDLRENYCRNPDGSESFWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
      :
DB      330 QYHEHDMTPENKCKDLRENYCRNPDGSESFWCFTTDPNIRVGYCSQIPNCDSHGQDC 389
QY      361 YRNGKNYVGNLSQTSRSGLTCSMWDKMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 420
      :
DB      390 YRNGKNYVGNLSQTSRSGLTCSMWDKMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 449
QY      421 CYTGNPLIPMDYCPISRCGEGDTTPTIV 447
      :
DB      450 CYTGNPLIPMDYCPISRCGEGDTTPTIV 476
```

```
RESULT 5
US-11-021-951-120
; Sequence 120, Application US/11021951
; Publication No. US20050175581A1
; GENERAL INFORMATION:
; APPLICANT: HAUPTS, Ulrich
; APPLICANT: KOLTERMANN, Andre
; APPLICANT: SCHEIDIG, Andreas
```

```
; APPLICANT: VOTSMERER, Christian
; APPLICANT: Ketting, Ulrich
; APPLICANT: COCO, Wayne Michael
; TITLE OF INVENTION: New Biological Entities And The Pharmaceutical
; FILE REFERENCE: 04156.000205
; CURRENT APPLICATION NUMBER: US/11/021,951
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: 10/872,198
; PRIOR FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 60/543,518
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/524,960
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: EP 04003058
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: EP 03025871
; PRIOR FILING DATE: 2003-11-11
; PRIOR APPLICATION NUMBER: EP 03025851
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: EP 03013819
; PRIOR FILING DATE: 2003-06-18
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-021-951-120
```

```
Query Match          99.9%; Score 2601; DB 20; Length 726;
Best Local Similarity 99.8%; Pred. No. 2e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ERRKRRYTHIEFKKSATTLIKIDPALKIKTKKVNADQCANRCTRNGKLPFTCKAFVFDK 60
      :
DB      30 QRRRRNTIHEFKKSATTLIKIDPALKIKTKKVNADQCANRCTRNGKLPFTCKAFVFDK 89
QY      61 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDYIRNCIIGKGSYKGTVAITKSGIKCOP 120
      :
DB      90 ARKQCLMPFPNSMSGVKEFGHEFDLYENKDYIRNCIIGKGSYKGTVAITKSGIKCOP 149
QY      121 WSSMIDHESFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNDEVRVEVCDIPQCESEVEC 180
      :
DB      150 WSSMIDHESFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNDEVRVEVCDIPQCESEVEC 209
QY      181 MTCNGSSYRGKDLMDHTSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRRPWC 240
      :
DB      210 MTCNGSSYRGKDLMDHTSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRRPWC 269
QY      241 YTLDPHTRMWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNGLPCQRMDS 300
      :
DB      270 YTLDPHTRMWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNGLPCQRMDS 329
QY      301 QYHEHDMTPENKCKDLRENYCRNPDGSESFWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
      :
DB      330 QYHEHDMTPENKCKDLRENYCRNPDGSESFWCFTTDPNIRVGYCSQIPNCDSHGQDC 389
QY      361 YRNGKNYVGNLSQTSRSGLTCSMWDKMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 420
      :
DB      390 YRNGKNYVGNLSQTSRSGLTCSMWDKMDLHRHIFWEPDASKLNNENYCRNPDGDAHGPW 449
QY      421 CYTGNPLIPMDYCPISRCGEGDTTPTIV 447
      :
DB      450 CYTGNPLIPMDYCPISRCGEGDTTPTIV 476
```

```
RESULT 6
US-08-605-221-2
; Sequence 2, Application US/08605221
; Publication No. US20030060403A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
```

TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
FILE REFERENCE: 2520-0101P
CURRENT APPLICATION NUMBER: US/08/605,221
CURRENT FILING DATE: 1996-03-15
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match 99.9%; Score 2601; DB 8; Length 728;
Best Local Similarity 99.8%; Pred. No. 2e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCITRNKGLPTCAFAVFDK 60
DB :
32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCITRNKGLPTCAFAVFDK 91
QY 61 ARKQCLWPFNSMSGVKGEGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKOP 120
DB :
92 ARKQCLWPFNSMSGVKGEGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKOP 151
QY 121 WSMIPIHEHSLPSSYRGKDIQENYCRNPGESEGPCFTSNPEVRYEVCIDIPQSEVYC 180
DB 152 WSMIPIHEHSLPSSYRGKDIQENYCRNPGESEGPCFTSNPEVRYEVCIDIPQSEVYC 211
QY 181 MTCNGESYRGIMDHTESGKICQRMHDQTPHRAKFLPERYPDKGDDNYCRNPDDAGHPW 240
DB 212 MTCNGESYRGIMDHTESGKICQRMHDQTPHRAKFLPERYPDKGDDNYCRNPDDAGHPW 271
QY 241 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQGGEGYRGVTNTIWMGIPCORWDS 300
DB 272 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQGGEGYRGVTNTIWMGIPCORWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDSSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDSSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 391
QY 361 YRGNGKNYMGSLQTRSGLTCSMDKNNMEDLHRAHFMEPDASKLNENYCRNPDDAGHPW 420
DB 392 YRGNGKNYMGSLQTRSGLTCSMDKNNMEDLHRAHFMEPDASKLNENYCRNPDDAGHPW 451
QY 421 CYTGNPLIPWDYCPISRCEGDTPTIIV 447
DB 452 CYTGNPLIPWDYCPISRCEGDTPTIIV 478

RESULT 7

US-10-813-805-10
Sequence 10, Application US/10813805
Publication No. US20040191902A1
GENERAL INFORMATION:
APPLICANT: PFIZER INC.
APPLICANT: Hamor, John E.
APPLICANT: Roach, Marsha L.
TITLE OF INVENTION: GROWTH AND DIFFERENTIATION OF STEM CELLS
FILE REFERENCE: PCS2028A
CURRENT APPLICATION NUMBER: US/10/813,805
CURRENT FILING DATE: 2004-03-30
PRIOR APPLICATION NUMBER: US 60/459,449
PRIOR FILING DATE: 2003-03-31
NUMBER OF SEQ ID NOS: 48
SOFTWARE: Patent In version 3.2
SEQ ID NO 10
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
US-10-813-805-10

Query Match 99.9%; Score 2601; DB 16; Length 728;

Best Local Similarity 99.8%; Pred. No. 2e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCITRNKGLPTCAFAVFDK 60
DB :
32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCITRNKGLPTCAFAVFDK 91
QY 61 ARKQCLWPFNSMSGVKGEGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKOP 120
DB 92 ARKQCLWPFNSMSGVKGEGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKOP 151
QY 121 WSMIPIHEHSLPSSYRGKDIQENYCRNPGESEGPCFTSNPEVRYEVCIDIPQSEVYC 180
DB 152 WSMIPIHEHSLPSSYRGKDIQENYCRNPGESEGPCFTSNPEVRYEVCIDIPQSEVYC 211
QY 181 MTCNGESYRGIMDHTESGKICQRMHDQTPHRAKFLPERYPDKGDDNYCRNPDDAGHPW 240
DB 212 MTCNGESYRGIMDHTESGKICQRMHDQTPHRAKFLPERYPDKGDDNYCRNPDDAGHPW 271
QY 241 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQGGEGYRGVTNTIWMGIPCORWDS 300
DB 272 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQGGEGYRGVTNTIWMGIPCORWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDSSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDSSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 391
QY 361 YRGNGKNYMGSLQTRSGLTCSMDKNNMEDLHRAHFMEPDASKLNENYCRNPDDAGHPW 420
DB 392 YRGNGKNYMGSLQTRSGLTCSMDKNNMEDLHRAHFMEPDASKLNENYCRNPDDAGHPW 451
QY 421 CYTGNPLIPWDYCPISRCEGDTPTIIV 447
DB 452 CYTGNPLIPWDYCPISRCEGDTPTIIV 478

RESULT 8

US-10-475-616-2
Sequence 2, Application US/10475616
Publication No. US20040236073A1
GENERAL INFORMATION:
APPLICANT: Gherardi, Ermanno
APPLICANT: Liehka, Daniel
APPLICANT: Blundell, Thomas L
APPLICANT: Chirgadze, Dmitry Y
TITLE OF INVENTION: The NK1 Fragment of Hepatocyte Growth Factor/Scatter
Factor (HGF/SF) and Variants Thereof, and Their Use
FILE REFERENCE: 2502489-991180
CURRENT APPLICATION NUMBER: US/10/475,616
CURRENT FILING DATE: 2003-10-22
PRIOR APPLICATION NUMBER: PCT/GB02/01941
PRIOR FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: GB 0110430.6
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
US-10-475-616-2

Query Match 99.9%; Score 2601; DB 16; Length 728;
Best Local Similarity 99.8%; Pred. No. 2e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCITRNKGLPTCAFAVFDK 60
DB :
32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCITRNKGLPTCAFAVFDK 91
QY 61 ARKQCLWPFNSMSGVKGEGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKOP 120
DB 92 ARKQCLWPFNSMSGVKGEGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKOP 151

QY 121 WSSMIFHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNBEVYEVCDIPQCSEVBC 180
DB 152 WSSMIFHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNBEVYEVCDIPQCSEVBC 211
QY 181 MTCNGSYRGKDLMDHTSSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRPWC 240
DB 212 MTCNGSYRGKDLMDHTSSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRPWC 271
QY 241 YTLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNGLPCQRMWS 300
DB 272 YTLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNGLPCQRMWS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRGNGKNYMGNTLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 420
DB 392 YRGNGKNYMGNTLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 451
QY 421 CYTGNDLIPWDYCPISRCBGDTTPTIV 447
DB 452 CYTGNDLIPWDYCPISRCBGDTTPTIV 478

RESULT 9

US-10-498-332-7
; Sequence 7, Application US/10498332
; Publication No. US20050113284A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Motoao
; APPLICANT: Higuchi, Toshio
; APPLICANT: Yamasaki, Yoshiaki
; APPLICANT: Orita, Takuya
; TITLE OF INVENTION: PHARMACEUTICAL AGENTS AND METHODS FOR
; FILE OF INVENTION: TISSUE AND VASCULAR REGENERATION
; FILE REFERENCE: 14539-016US1
; CURRENT APPLICATION NUMBER: US/10/498,332
; CURRENT FILING DATE: 2004-06-11
; PRIOR APPLICATION NUMBER: PCT/JP02/13014
; PRIOR FILING DATE: 2002-12-12
; PRIOR APPLICATION NUMBER: JP 2002-352924
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: JP 2001-380158
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-498-332-7

Query Match 99.9%; Score 2601; DB 17; Length 728;
Best Local Similarity 99.8%; Pred. No. 2e-211;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKRRTTIEFKKSATTTLIKIDPALKITKKNVTADQCANRCTRNKGLPFTCKAVFPA 60
DB 32 ORRRRTTIEFKKSATTTLIKIDPALKITKKNVTADQCANRCTRNKGLPFTCKAVFPA 91
QY 61 ARKQCLMPFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120
DB 92 ARKQCLMPFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151
QY 121 WSSMIFHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNBEVYEVCDIPQCSEVBC 180
DB 152 WSSMIFHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNBEVYEVCDIPQCSEVBC 211
QY 181 MTCNGSYRGKDLMDHTSSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRPWC 240
DB 212 MTCNGSYRGKDLMDHTSSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRPWC 271

QY 241 YTLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNGLPCQRMWS 300
DB 272 YTLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNGLPCQRMWS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRGNGKNYMGNTLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 420
DB 392 YRGNGKNYMGNTLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 451
QY 421 CYTGNDLIPWDYCPISRCBGDTTPTIV 447
DB 452 CYTGNDLIPWDYCPISRCBGDTTPTIV 478

RESULT 10

US-10-149-103A-1
; Sequence 1, Application US/10149103A
; Publication No. US20040121945A1
; GENERAL INFORMATION:
; APPLICANT: Liang, Hong
; APPLICANT: Sim, Kim Lee
; APPLICANT: Day, Clara
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0592 43170-274540
; CURRENT APPLICATION NUMBER: US/10/149,103A
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/US00/34039
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-149-103A-1

Query Match 98.5%; Score 2565; DB 16; Length 727;
Best Local Similarity 98.2%; Pred. No. 2.3e-208;
Matches 438; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 2 RKRRTTIEFKKSATTTLIKIDPALKITKKNVTADQCANRCTRNKGLPFTCKAVFPA 61
DB 33 KRRRTTIEFKKSATTTLIKIDPALKITKKNVTADQCANRCTRNKGLPFTCKAVFPA 92
QY 62 RKQCLMPFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 121
DB 93 RKQCLMPFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 152
QY 122 SSMIFHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNBEVYEVCDIPQCSEVBC 181
DB 153 SSMIFHEHSFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNBEVYEVCDIPQCSEVBC 212
QY 182 TONGSYRGKDLMDHTSSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRPWC 241
DB 213 TONGSYRGKDLMDHTSSGKICQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQRPWC 272
QY 242 YTLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNGLPCQRMWS 301
DB 273 YTLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNGLPCQRMWS 332
QY 302 YHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 361
DB 333 YHKHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 392
QY 362 RGNKKNYMGNTLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 421
DB 393 RGNKKNYMGNTLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 452
QY 422 YTGNDLIPWDYCPISRCBGDTTPTIV 447

COUNTRY: UNITED STATES OF AMERICA
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/5 inch diskette
COMPUTER: PC'S LIMITED SYSTEM 200
OPERATING SYSTEM: WORD FOR WINDOWS
SOFTWARE: MICROSOFT WINDOWS 98
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/133,912
FILING DATE: 25-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/700,519
FILING DATE: 26-Aug-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mayne, Milton J.
REGISTRATION NUMBER: 17,906
REFERENCE/DOCKET NUMBER: U-WP-5240PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-683-8150
TELEFAX: 212-532-4285
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 723
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-133-912-19
Query Match 97.5%; Score 2539.5; DB 13; Length 723;
Best Local Similarity 97.8%; Pred. No. 3.2e-206;
Matches 437; Conservative 1; Mismatches 4; Indels 5; Gaps 1;
QY 1 EKKRRNTHEFKRSKATTLIKIDPALKIKTKKVNADOCANRCTRNKGLPTCKAFVFDK 60
DB 32 QKRRNTTIEFKRSATTLIKIDPALAIATAVAVNTADOCANRCTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIGKRSYKGTVISITKSGIKCP 120
DB 92 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIGKRSYKGTVISITKSGIKCP 151
QY 121 WSSMTI PHEHSFLPSYRGKLOENYCRNPRGEGGFWCTSNPEVRYEYCDIPQCEVEEC 180
DB 152 WSSMTI PHEH-----SYRGKLOENYCRNPRGEGGFWCTSNPEVRYEYCDIPQCEVEEC 206
QY 181 MTCNSESYSYGLMDHTESGKICQRWDHQTBRHKFLPERYPDKGPDNVCNRPDQAPRWC 240
DB 207 MTCNSESYSYGLMDHTESGKICQRWDHQTBRHKFLPERYPDKGPDNVCNRPDQAPRWC 266
QY 241 YTLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGTVNTIWMGI PCQRWDS 300
DB 267 YTLDPHTRWYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGTVNTIWMGI PCQRWDS 326
QY 301 QYRHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVYCSQIPNCDSHGQDC 360
DB 327 QYRHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVYCSQIPNCDSHGQDC 386
QY 361 YRNGNGNVMGNLSQTRSGLTCSMMDKNMMDLHRHIFWEPDASKLNENVCNRPDDAHGPM 420
DB 387 YRNGNGNVMGNLSQTRSGLTCSMMDKNMMDLHRHIFWEPDASKLNENVCNRPDDAHGPM 446
QY 421 CYTGNPLIPMDYCPISRCGDTTPTIV 447
DB 447 CYTGNPLIPMDYCPISRCGDTTPTIV 473

Search completed: October 3, 2005, 07:42:28
Job time : 243.355 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:32 ; Search time 1375.69 Seconds
(without alignments)
24.255 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604
Sequence: 1 EKKRRNTIHEFKSAKTLTI.....IPWDYCPISRCBGDTPTTIV 447

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2601	99.9	697	2	US-08-460-890A-50 Sequence 50, Appl
2	2601	99.9	697	3	US-08-167-641C-50 Sequence 50, Appl
3	2601	99.9	697	3	US-08-460-971A-50 Sequence 50, Appl
4	2601	99.9	697	3	US-08-462-040-50 Sequence 50, Appl
5	2601	99.9	728	1	US-07-815-333A-2 Sequence 2, Appl
6	2601	99.9	728	4	US-08-087-783A-22 Sequence 2, Appl
7	2601	99.9	728	4	US-08-605-221-2 Sequence 2, Appl
8	2563.5	98.4	487	4	US-09-601-040A-27 Sequence 27, Appl
9	2563.5	98.4	717	4	US-09-601-040A-6 Sequence 8, Appl
10	2563.5	98.4	717	4	US-09-601-040A-8 Sequence 6, Appl
11	2563.5	98.4	723	1	US-07-838-410-1 Sequence 1, Appl
12	2563.5	98.4	723	1	US-08-290-937B-1 Sequence 1, Appl
13	2563.5	98.4	723	1	US-08-290-937B-2 Sequence 1, Appl
14	2563.5	98.4	723	1	US-08-404-643-1 Sequence 1, Appl
15	2563.5	98.4	723	1	US-09-194-326-1 Sequence 1, Appl
16	2563.5	98.4	723	4	US-09-600-991-18 Sequence 18, Appl
17	2563.5	98.4	723	4	US-08-605-221-4 Sequence 4, Appl
18	2563.5	98.4	723	4	US-09-601-040A-10 Sequence 10, Appl
19	2563.5	98.4	723	4	US-09-601-040A-2 Sequence 4, Appl
20	2563.5	98.4	729	4	US-09-601-040A-4 Sequence 4, Appl
21	2560.5	98.3	723	1	US-08-030-410-3 Sequence 3, Appl
22	2557.5	98.2	723	1	US-08-290-937B-3 Sequence 3, Appl
23	2541.5	97.6	723	3	US-09-194-326-2 Sequence 2, Appl
24	2541.5	97.6	723	3	US-08-700-519J-18 Sequence 18, Appl
25	2539.5	97.5	723	3	US-09-194-326-3 Sequence 3, Appl
26	2539.5	97.5	723	3	US-08-700-519J-19 Sequence 19, Appl
27	1780.5	68.4	574	4	US-09-600-991-2 Sequence 2, Appl

28	1471	56.5	290	4	US-08-484-841A-8 Sequence 8, Appl
29	1447.5	55.6	563	4	US-09-600-991-4 Sequence 4, Appl
30	1180.5	45.3	711	1	US-08-184-012C-8 Sequence 8, Appl
31	1180.5	45.3	711	1	US-08-334-177-2 Sequence 2, Appl
32	1180.5	45.3	711	2	US-08-666-082B-1 Sequence 1, Appl
33	1180.5	45.3	711	2	US-08-766-982-2 Sequence 2, Appl
34	1180.5	45.3	711	3	US-09-296-219-2 Sequence 2, Appl
35	1180.5	45.3	711	4	US-09-600-991-20 Sequence 20, Appl
36	1180.5	45.3	711	4	US-09-601-040A-12 Sequence 12, Appl
37	1180.5	45.3	711	4	US-09-949-016-6981 Sequence 6981, Ap
38	1180.5	45.3	711	5	PCT-US95-13830-2 Sequence 2, Appl
39	1125	43.2	716	2	US-08-766-982-1 Sequence 1, Appl
40	1125	43.2	716	3	US-09-296-219-1 Sequence 1, Appl
41	1066	40.9	790	4	US-08-991-761A-13 Sequence 13, Appl
42	1059	40.7	812	1	US-08-248-629A-1 Sequence 1, Appl
43	1059	40.7	812	1	US-08-451-932-1 Sequence 1, Appl
44	1059	40.7	812	1	US-08-452-260-1 Sequence 1, Appl
45	1059	40.7	812	1	US-08-326-785-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-460-890A-50
; Sequence 50, Application US/08460890A
; Patent No. 5994109
; GENERAL INFORMATION:
; APPLICANT: Moo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OR INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,890A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-890A-50

Query Match
Best Local Similarity 99.9%; Score 2601; DB 2; Length 697;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRRNTIHEPKSAKTTLIKIDPALKIKYKVNADQCANRCTRNKGLPTCKAFVFDK 60
DB 1 QKRRNTIHEPKSAKTTLIKIDPALKIKYKVNADQCANRCTRNKGLPTCKAFVFDK 60
QY 61 ARKQCLMPFPNMSGKKEFGHEFDLYENKDYINNCIIIGRSYKGVSIITKSGIKOP 120
DB 61 ARKQCLMPFPNMSGKKEFGHEFDLYENKDYINNCIIIGRSYKGVSIITKSGIKOP 120
QY 121 WSSMTPEHSHFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCIDIPQSEVYC 180
DB 121 WSSMTPEHSHFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCIDIPQSEVYC 180
QY 181 MTCNGESYRGLMDHTESGKICQRMHDQTPHKKFLPERYPDKGFPDNNYCRNPDGQRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQRMHDQTPHKKFLPERYPDKGFPDNNYCRNPDGQRPWC 240
QY 241 YLDPHTRWECALIKTCADNTMNDTVPLETTTCIOGGEGYRGVNTIWMGIPQQRWDS 300
DB 241 YLDPHTRWECALIKTCADNTMNDTVPLETTTCIOGGEGYRGVNTIWMGIPQQRWDS 300
QY 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
DB 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
QY 361 YRGNKNYMGNISQTRSGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDHAGPW 420
DB 361 YRGNKNYMGNISQTRSGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDHAGPW 420
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447

RESULT 2
US-08-167-641C-50
Sequence 50, Application US/08167641C
Patent No. 6033884

GENERAL INFORMATION:
APPLICANT: MCO, Savio L. C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,641C
FILING DATE: December 14, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/855,389

FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 205/012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-167-641C-50

Query Match
Best Local Similarity 99.9%; Score 2601; DB 3; Length 697;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRRNTIHEPKSAKTTLIKIDPALKIKYKVNADQCANRCTRNKGLPTCKAFVFDK 60
DB 1 QKRRNTIHEPKSAKTTLIKIDPALKIKYKVNADQCANRCTRNKGLPTCKAFVFDK 60
QY 61 ARKQCLMPFPNMSGKKEFGHEFDLYENKDYINNCIIIGRSYKGVSIITKSGIKOP 120
DB 61 ARKQCLMPFPNMSGKKEFGHEFDLYENKDYINNCIIIGRSYKGVSIITKSGIKOP 120
QY 121 WSSMTPEHSHFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCIDIPQSEVYC 180
DB 121 WSSMTPEHSHFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVRYEVCIDIPQSEVYC 180
QY 181 MTCNGESYRGLMDHTESGKICQRMHDQTPHKKFLPERYPDKGFPDNNYCRNPDGQRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQRMHDQTPHKKFLPERYPDKGFPDNNYCRNPDGQRPWC 240
QY 241 YLDPHTRWECALIKTCADNTMNDTVPLETTTCIOGGEGYRGVNTIWMGIPQQRWDS 300
DB 241 YLDPHTRWECALIKTCADNTMNDTVPLETTTCIOGGEGYRGVNTIWMGIPQQRWDS 300
QY 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
DB 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
QY 361 YRGNKNYMGNISQTRSGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDHAGPW 420
DB 361 YRGNKNYMGNISQTRSGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDHAGPW 420
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447

RESULT 3
US-08-460-971A-50
Sequence 50, Application US/08460971A
Patent No. 6150168

GENERAL INFORMATION:
APPLICANT: MCO, Savio L. C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,971A
FILING DATE: December 14, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/855,389


```

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,971A
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-971A-50

```

Query Match 99.9%; Score 2601; DB 3; Length 697;

Best Local Similarity 99.8%; Pred. No. 2.7e-214; Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ERKRNTIHEPKSAKTLIKIDPALKIKTKKVNADOCANRTRNGKLPFTCKAFVFDK 60
DB 1 QKRKRNTIHEPKSAKTLIKIDPALKIKTKKVNADOCANRTRNGKLPFTCKAFVFDK 60
QY 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKOP 120
DB 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKOP 120
QY 121 WSSMIPIHESFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDDIPQSEVBC 180
DB 121 WSSMIPIHESFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDDIPQSEVBC 180
QY 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHRRKFLBERYPDKGFDNNYCRNPDGQRPWC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHRRKFLBERYPDKGFDNNYCRNPDGQRPWC 240
QY 241 YTLDPHTTRREYCAIKTCADNTAMDTPVPLETTECIGQSGRGYNTITWNGIPQCRWMS 300
DB 241 YTLDPHTTRREYCAIKTCADNTAMDTPVPLETTECIGQSGRGYNTITWNGIPQCRWMS 300
QY 301 QYPIHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRYVCSQIPNCDSHGQDC 360
DB 301 QYPIHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRYVCSQIPNCDSHGQDC 360
QY 361 YRNGNGXNYGNLSQTSRGLTCSWMDXNMDLHRHIFWEPDASKLNNENYCRNPDDAHGPW 420
DB 361 YRNGNGXNYGNLSQTSRGLTCSWMDXNMDLHRHIFWEPDASKLNNENYCRNPDDAHGPW 420
QY 421 YTGNGXNYGNLSQTSRGLTCSWMDXNMDLHRHIFWEPDASKLNNENYCRNPDDAHGPW 420
DB 421 YTGNGXNYGNLSQTSRGLTCSWMDXNMDLHRHIFWEPDASKLNNENYCRNPDDAHGPW 420
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

```

RESULT 4

US-08-462-040-50

Sequence 50, Application US/08462040

Patent No. 6177554

GENERAL INFORMATION:

APPLICANT: Woo, Savio L.C.

APPLICANT: Smith, Louis C.

APPLICANT: Cristiano, Richard J.

APPLICANT: Gottchalk, Stephen

TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND

TITLE OF INVENTION: METHODS OF USE

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,040

FILING DATE: June 5, 1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/167,641

FILING DATE: December 14, 1993

APPLICATION NUMBER: 07/855,389

FILING DATE: March 20, 1992

APPLICATION NUMBER: PCT/US93/02725

FILING DATE: March 19, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 212/078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ. ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 697 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-462-040-50

Query Match 99.9%; Score 2601; DB 3; Length 697;

Best Local Similarity 99.8%; Pred. No. 2.7e-214; Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ERKRNTIHEPKSAKTLIKIDPALKIKTKKVNADOCANRTRNGKLPFTCKAFVFDK 60
DB 1 QKRKRNTIHEPKSAKTLIKIDPALKIKTKKVNADOCANRTRNGKLPFTCKAFVFDK 60
QY 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKOP 120
DB 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKOP 120
QY 121 WSSMIPIHESFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDDIPQSEVBC 180
DB 121 WSSMIPIHESFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDDIPQSEVBC 180
QY 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHRRKFLBERYPDKGFDNNYCRNPDGQRPWC 240
DB 181 MTCNGESYRGKDLMDHTESGKICQRMHDQTPHRRKFLBERYPDKGFDNNYCRNPDGQRPWC 240

```

Db 181 MTCNGESYRGMLMDHTESGKICQRMHDQTPRHKKFLPERYPDKGPDNCRNPDGQPRPWC 240
Qy 241 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTWNTMNGIPQGRWDS 300
Db 241 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTWNTMNGIPQGRWDS 300
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFITDPNIRVGYSQIIPNCDSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFITDPNIRVGYSQIIPNCDSHGQDC 360
Qy 361 YRGNGKNYMGSLQSTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNPDDAHGPW 420
Db 361 YRGNGKNYMGSLQSTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNPDDAHGPW 420
Qy 421 CYTGNPLIPMDYCPISRCGEGDTTPTIV 447
Db 421 CYTGNPLIPMDYCPISRCGEGDTTPTIV 447

RESULT 5

US-07-815-333A-2
; Sequence 2, Application US/07815333A
; Patent No. 5342831

; GENERAL INFORMATION:
; APPLICANT: Nakamura, Toshikazu
; APPLICANT: Matsumoto, Kunio
; TITLE OF INVENTION: EPITHELIOCYTE GROWTH ACCELERATOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/815.333A
; FILING DATE: 19911227
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Robert F.
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 44069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-815-333A-2

Query Match 99.9%; Score 2601; DB 1; Length 728;
Best Local Similarity 99.8%; Pred. No. 2.9e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKRKRNTHFEFKSAKTTLIKIDPLAKITKKNVNTADQCANCTFNKGLPTCKAFVFDK 60
Db 32 QKRNKNTTHFEFKSAKTTLIKIDPLAKITKKNVNTADQCANCTFNKGLPTCKAFVFDK 91
Qy 61 ARKQCLWFPNMSGGVKKEFGHEFDLVENKDYINNCIIIGKRSYKGTWISIKSGIKOP 120
Db 92 ARKQCLWFPNMSGGVKKEFGHEFDLVENKDYINNCIIIGKRSYKGTWISIKSGIKOP 151

Qy 121 WSMIIPHEHSFLBSSYRGKDLQENYCRNPRGEGCPWCFITSNPEVRYEVCIDIPQCSVEEC 180
Db 152 WSMIIPHEHSFLBSSYRGKDLQENYCRNPRGEGCPWCFITSNPEVRYEVCIDIPQCSVEEC 211
Qy 181 MTCNGESYRGMLMDHTESGKICQRMHDQTPRHKKFLPERYPDKGPDNCRNPDGQPRPWC 240
Db 212 MTCNGESYRGMLMDHTESGKICQRMHDQTPRHKKFLPERYPDKGPDNCRNPDGQPRPWC 271
Qy 241 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTWNTMNGIPQGRWDS 300
Db 272 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTWNTMNGIPQGRWDS 331
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFITDPNIRVGYSQIIPNCDSHGQDC 360
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFITDPNIRVGYSQIIPNCDSHGQDC 391
Qy 361 YRGNGKNYMGSLQSTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNPDDAHGPW 420
Db 392 YRGNGKNYMGSLQSTRSGLTCSMDKXNMDLHRHIFWEPDASKLNENYCRNPDDAHGPW 451
Qy 421 CYTGNPLIPMDYCPISRCGEGDTTPTIV 447
Db 452 CYTGNPLIPMDYCPISRCGEGDTTPTIV 478

RESULT 6

US-08-087-783A-22
; Sequence 22, Application US/08087783A
; Patent No. 5547856

; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087.783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: Amino Acid
; TOPOLOGY: linear
; US-08-087-783A-22

Query Match 99.9%; Score 2601; DB 1; Length 728;
Best Local Similarity 99.8%; Pred. No. 2.9e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ERKRRTIHEFKSAKTLIKIDPAKIKTKKVNADOCANRCTRNNKGLPFTCKAFVFDK 60
Db 32 QRRKRRTIHEFKSAKTLIKIDPAKIKTKKVNADOCANRCTRNNKGLPFTCKAFVFDK 91
QY 61 ARKQCLMFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITSGIKCOP 120
Db 92 ARKQCLMFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITSGIKCOP 151
QY 121 WSSMIPIHEHSFLPSSYRGKDLQENYCNPRGEGGPGCFTSNPEVRYEVCDDIPQCSBVEVC 180
Db 152 WSSMIPIHEHSFLPSSYRGKDLQENYCNPRGEGGPGCFTSNPEVRYEVCDDIPQCSBVEVC 211
QY 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQPRIHKKFLPERYPDKGFDNNYCNPNPQGPAPWC 240
Db 212 MTCNGESYRGKDLMDHTESGKICQRMWDHQPRIHKKFLPERYPDKGFDNNYCNPNPQGPAPWC 271
QY 241 YTLDPHTRWYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTNTIINGIIPCORWDS 300
Db 272 YTLDPHTRWYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTNTIINGIIPCORWDS 331
QY 301 QYHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 332 QYHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRGNGKNVNGNLQOTSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCNPPDDAHGPM 420
Db 392 YRGNGKNVNGNLQOTSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCNPPDDAHGPM 451
QY 421 CYTGNPLIPWDYCPISRCGDDTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGDDTPTIV 478

RESULT 7
US-08-605-221-2
; Sequence 2, Application US/08605221
; Patent No. 6699837
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match 99.9%; Score 2601; DB 4; Length 728;
Best Local Similarity 99.8%; Pred. No. 2.9e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTIHEFKSAKTLIKIDPAKIKTKKVNADOCANRCTRNNKGLPFTCKAFVFDK 60
Db 32 QRRKRRTIHEFKSAKTLIKIDPAKIKTKKVNADOCANRCTRNNKGLPFTCKAFVFDK 91
QY 61 ARKQCLMFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITSGIKCOP 120
Db 92 ARKQCLMFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITSGIKCOP 151
QY 121 WSSMIPIHEHSFLPSSYRGKDLQENYCNPRGEGGPGCFTSNPEVRYEVCDDIPQCSBVEVC 180
Db 152 WSSMIPIHEHSFLPSSYRGKDLQENYCNPRGEGGPGCFTSNPEVRYEVCDDIPQCSBVEVC 211
QY 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQPRIHKKFLPERYPDKGFDNNYCNPNPQGPAPWC 240
Db 212 MTCNGESYRGKDLMDHTESGKICQRMWDHQPRIHKKFLPERYPDKGFDNNYCNPNPQGPAPWC 271
```

```
QY 241 YTLDPHTRWYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTNTIINGIIPCORWDS 300
Db 272 YTLDPHTRWYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTNTIINGIIPCORWDS 331
QY 301 QYHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 332 QYHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 361 YRGNGKNVNGNLQOTSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCNPPDDAHGPM 420
Db 392 YRGNGKNVNGNLQOTSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCNPPDDAHGPM 451
QY 421 CYTGNPLIPWDYCPISRCGDDTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGDDTPTIV 478

RESULT 8
US-09-601-040A-27
; Sequence 27, Application US/09601040A
; Patent No. 6730657
; GENERAL INFORMATION:
; APPLICANT: MEDICO, Enzo et al.
; TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
; FILE REFERENCE: 0471-0161P
; CURRENT APPLICATION NUMBER: US/09/601,040A
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-601-040A-27

Query Match 98.4%; Score 2563.5; DB 4; Length 487;
Best Local Similarity 98.7%; Pred. No. 2.8e-211;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRRTIHEFKSAKTLIKIDPAKIKTKKVNADOCANRCTRNNKGLPFTCKAFVFDK 60
Db 32 QRRKRRTIHEFKSAKTLIKIDPAKIKTKKVNADOCANRCTRNNKGLPFTCKAFVFDK 91
QY 61 ARKQCLMFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITSGIKCOP 120
Db 92 ARKQCLMFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITSGIKCOP 151
QY 121 WSSMIPIHEHSFLPSSYRGKDLQENYCNPRGEGGPGCFTSNPEVRYEVCDDIPQCSBVEVC 180
Db 152 WSSMIPIHEH-----SYRGKDLQENYCNPRGEGGPGCFTSNPEVRYEVCDDIPQCSBVEVC 206
QY 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQPRIHKKFLPERYPDKGFDNNYCNPNPQGPAPWC 240
Db 207 MTCNGESYRGKDLMDHTESGKICQRMWDHQPRIHKKFLPERYPDKGFDNNYCNPNPQGPAPWC 266
QY 241 YTLDPHTRWYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTNTIINGIIPCORWDS 300
Db 267 YTLDPHTRWYCAIKTCADNTMNDTVPLETTTECICQGGEGYRGVTNTIINGIIPCORWDS 326
QY 301 QYHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
Db 327 QYHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 386
QY 361 YRGNGKNVNGNLQOTSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCNPPDDAHGPM 420
Db 387 YRGNGKNVNGNLQOTSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCNPPDDAHGPM 446
QY 421 CYTGNPLIPWDYCPISRCGDDTPTIV 447
Db 447 CYTGNPLIPWDYCPISRCGDDTPTIV 473

RESULT 9
```


APPLICATION NUMBER: PCT/JP91/00942
FILING DATE: 15-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 185852-1990
FILING DATE: 13-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN L.
REGISTRATION NUMBER: 17,746
REFERENCE/DOCKET NUMBER: WAK 110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
TELEX: 64191
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Fibroblast
FEATURE:
NAME/KEY: Domain
LOCATION: 393..405 /note= "INTERNAL AMINO ACID
OTHER INFORMATION: SEQUENCE IN ALPHA-CHAIN"
FEATURE:
NAME/KEY: Protein
LOCATION: 490..505 /note= "N-TERMINAL AMINO ACID
OTHER INFORMATION: SEQUENCE OF BETA-CHAIN"
FEATURE:
NAME/KEY: Domain
LOCATION: 605..623 /note= "INTERNAL AMINO ACID
OTHER INFORMATION: SEQUENCE IN BETA-CHAIN"
US-07-838-410-1

Query Match 98.4%; Score 2563.5; DB 1; Length 723;
Best Local Similarity 98.7%; Pred. No. 4.7e-211;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 EKKRRNTIEFKSAKTLTIKIDPALKITKKVNTADQCANCTRNKGLPFTCKAVFPK 60
:|||||
DB 32 QKRRNTIEFKSAKTLTIKIDPALKITKKVNTADQCANCTRNKGLPFTCKAVFPK 91
:|||||

QY 61 ARKQCLMPFPNMSGVKKFEGHEPDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 120
:|||||
DB 92 ARKQCLMPFPNMSGVKKFEGHEPDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 151
:|||||

QY 121 WSSMTPIHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180
:|||||
DB 152 WSSMTPIHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 206
:|||||

QY 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQPFRHKFLPERYPDKGFDNRYCRNPDGQRPWC 240
:|||||
DB 207 MTCNGESYRGKDLMDHTESGKICQRMWDHQPFRHKFLPERYPDKGFDNRYCRNPDGQRPWC 266
:|||||

QY 241 YLIDPTHTWEYCAIKTCADNTMNDIVPLETTECIGQGEGRYGVNTIWNIGIPQGRWDS 300
:|||||
DB 267 YLIDPTHTWEYCAIKTCADNTMNDIVPLETTECIGQGEGRYGVNTIWNIGIPQGRWDS 326
:|||||

QY 301 QYRPHEDMTPEKFKCDLRENYCRNPDGSESPWCFTTDNIRYGVCSQIPNCMSHGQPC 360
:|||||
DB 327 QYRPHEDMTPEKFKCDLRENYCRNPDGSESPWCFTTDNIRYGVCSQIPNCMSHGQPC 386
:|||||

QY 361 YRNGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASKLANENYCRNPDGDAHGPW 420
:|||||
DB 387 YRNGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASKLANENYCRNPDGDAHGPW 446
:|||||

QY 421 CYTGNPLIMWDCPIRSCGDTTPTIV 447

|||||
DB 447 CYTGNPLIMWDCPIRSCGDTTPTIV 473

RESULT 12
US-08-290-937B-1
Sequence 1, Application US/08290937B
Patent No. 5648233
GENERAL INFORMATION:
APPLICANT: YAMAGUCHI, KYOJI
APPLICANT: SHIMA, NOBUYUKI
APPLICANT: MURAKAMI, AKIHIKO
APPLICANT: GOTO, MASAOKI
APPLICANT: TSUDA, EISUKE
APPLICANT: MASUNAGA, HIROAKI
APPLICANT: TAKAHARA, REIKO
APPLICANT: OOGAKI, RUMIKO
APPLICANT: UEDA, MASATSUGU
APPLICANT: HIGASHIO, KANJI
TITLE OF INVENTION: MODIFIED TCF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSER: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,937B
FILING DATE: 19-AUG-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-022
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-937B-1

Query Match 98.4%; Score 2563.5; DB 1; Length 723;
Best Local Similarity 98.7%; Pred. No. 4.7e-211;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 EKKRRNTIEFKSAKTLTIKIDPALKITKKVNTADQCANCTRNKGLPFTCKAVFPK 60
:|||||
DB 32 QKRRNTIEFKSAKTLTIKIDPALKITKKVNTADQCANCTRNKGLPFTCKAVFPK 91
:|||||

QY 61 ARKQCLMPFPNMSGVKKFEGHEPDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 120
:|||||
DB 92 ARKQCLMPFPNMSGVKKFEGHEPDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 151
:|||||

QY 121 WSSMTPIHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180
:|||||
DB 152 WSSMTPIHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 206
:|||||

QY 181 MTCNGESYRGKDLMDHTESGKICQRMWDHQPFRHKFLPERYPDKGFDNRYCRNPDGQRPWC 240
:|||||
DB 207 MTCNGESYRGKDLMDHTESGKICQRMWDHQPFRHKFLPERYPDKGFDNRYCRNPDGQRPWC 266
:|||||

QY 241 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQOGGEGYRGTVNTIWMGIPCORWDS 300
DB 267 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQOGGEGYRGTVNTIWMGIPCORWDS 326
QY 301 QYPHEHDMTPENFKCKDIRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 327 QYPHEHDMTPENFKCKDIRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 386
QY 361 YRGNGKNYMGNLSTQSRGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDGDAHGPM 420
DB 387 YRGNGKNYMGNLSTQSRGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDGDAHGPM 446
QY 421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447
DB 447 CYTGNPLIPWDYCPISRCEGDTTPTIV 473

RESULT 13
US-08-290-937B-2
; Sequence 2, Application US/08290937B
; Patent No. 5648233
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, KYOJI
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: MURAKAMI, AKIHIKO
; APPLICANT: GOTO, MASAKI
; APPLICANT: TSUDA, EISUKE
; APPLICANT: MASUNAGA, HIROAKI
; APPLICANT: TAKAHIRA, REIKO
; APPLICANT: OOGAKI, FUMIKO
; APPLICANT: UEDA, MASATSUGU
; APPLICANT: HIGASHIO, KANUJI
; TITLE OF INVENTION: MODIFIED TCF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,937B
; FILING DATE: 19-AUG-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-290-937B-2

Query Match 98.4%; Score 2563.5; DB 1; Length 723;
Best Local Similarity 98.7%; Pred. No. 4.7e-211;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRRNTIEFKKSATYTIKIDPAKIKTKKVNNTADQCANCTRNKGLPFTCKAFVFDK 60
DB 32 QKRRNTIEFKKSATYTIKIDPAKIKTKKVNNTADQCANCTRNKGLPFTCKAFVFDK 91

QY 61 ARKOLWFENSMSSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTITKSIKQOP 120
DB 92 ARKOLWFENSMSSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTITKSIKQOP 151
QY 121 WSMIIPHEHSLPSSYRKQDIQENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 180
DB 152 WSMIIPHEHSLPSSYRKQDIQENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 206
QY 181 MTCNGESYRGIMDHTESGKIQRMWDHQPHEHKLPERYPKGFDPNYCRNPDGPRWC 240
DB 207 MTCNGESYRGIMDHTESGKIQRMWDHQPHEHKLPERYPKGFDPNYCRNPDGPRWC 266
QY 241 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQOGGEGYRGTVNTIWMGIPCORWDS 300
DB 267 YTLDPHTRMEYCAIKTCADNTMNDTVDVPLETTECIQOGGEGYRGTVNTIWMGIPCORWDS 326
QY 301 QYPHEHDMTPENFKCKDIRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 360
DB 327 QYPHEHDMTPENFKCKDIRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 386
QY 361 YRGNGKNYMGNLSTQSRGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDGDAHGPM 420
DB 387 YRGNGKNYMGNLSTQSRGLTCSMDKXNEDLHRHIFWEPDASKLNENYCRNPDGDAHGPM 446
QY 421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447
DB 447 CYTGNPLIPWDYCPISRCEGDTTPTIV 473

RESULT 14
US-08-404-643-1
; Sequence 1, Application US/08404643
; Patent No. 5658742
; GENERAL INFORMATION:
; APPLICANT: HIGASHIO, KANUJI
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: OOGAKI, FUMIKO
; TITLE OF INVENTION: MONOCLONAL ANTIBODY
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,643
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-034 (3999/35)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-404-643-1

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:29 ; Search time 39.775 Seconds
(without alignments)
1069.209 Million cell updates/sec

Title: US-09-674-377B-2

Perfect score: 2579
Sequence: 1 EKKRRNTIHEFKSAKTLI.....IPWDYCPISRCGDTPTIV 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2563.5	99.4	728	1 JH0579	hepatocyte growth
2	2370.5	91.9	728	1 A60185	hepatocyte growth
3	2263.5	91.6	728	1 A35644	hepatocyte growth
4	1830	71.0	710	1 I51283	hepatocyte growth
5	1750	67.9	411	2 I51285	hepatocyte growth
6	1171	45.4	711	1 A47136	macrophage-stimula
7	1133.5	44.0	716	1 JCS061	macrophage-stimula
8	1114.5	43.2	716	1 A40332	macrophage-stimula
9	1054.5	40.9	790	1 PLRG	plasma (EC 3.4.21
10	1039.5	40.3	812	1 PLMS	plasma (EC 3.4.21
11	1036	40.2	810	2 I46260	plasma (EC 3.4.21
12	1021	39.6	810	2 B30848	plasma (EC 3.4.21
13	1019.5	39.5	812	2 B30848	plasma (EC 3.4.21
14	1018.5	39.5	810	1 PLBO	plasma (EC 3.4.21
15	927	35.9	2869	2 T1818	apolioprotein(a)
16	864.5	33.5	4548	1 S00657	apolioprotein(a) (EC
17	840	32.6	1420	2 A32869	apolioprotein(a)
18	457	17.7	460	2 B61545	plasma (EC 3.4.21
19	450.5	17.5	455	2 A61545	plasma (EC 3.4.21
20	406	15.7	336	2 S33879	plasma precursor
21	357.5	13.9	625	1 TBBO	chrombin (EC 3.4.2
22	353	13.7	169	2 A40522	plasma (EC 3.4.21
23	352.5	13.6	618	2 A35827	chrombin (EC 3.4.2
24	350	13.3	617	1 TBHU	chrombin (EC 3.4.2
25	344	13.3	617	1 S10511	chrombin (EC 3.4.2
26	302.5	11.7	562	1 UKHUT	c-plasminogen acti
27	302	11.7	559	1 A35029	c-plasminogen acti
28	291	11.3	559	1 A29941	c-plasminogen acti
29	270	10.5	120	2 E61545	plasma (EC 3.4.21

30	268	10.4	123	2 C61545	plasma (EC 3.4.21
31	260.5	10.1	89	2 A60140	plasma (EC 3.4.21
32	237.5	9.2	946	1 A47299	ror-related recept
33	236.5	9.2	943	2 B45082	neurotrophic recep
34	228.5	8.9	603	2 S28941	coagulation factor
35	225	8.7	291	2 I38098	c-plasminogen acti
36	213.5	8.3	937	2 A45082	neurotrophic recep
37	204	7.9	655	1 A46688	hepatocyte growth
38	200	7.8	615	1 KFHU12	coagulation factor
39	192.5	7.5	558	2 JCS078	plasma hyaluronan-
40	189	7.3	593	2 S45281	coagulation factor
41	186.5	7.2	560	1 UC4795	plasma hyaluronan-
42	182.5	7.1	806	2 T18840	hypothetical prote
43	158.5	6.1	433	1 JN0560	u-plasminogen acti
44	158.5	6.1	685	1 A48289	neurotrophic recep
45	158	6.1	442	1 UKPG	u-plasminogen acti

ALIGNMENTS

RESULT 1

JH0579
hepatocyte growth factor precursor [validated] - human
N/Alternate names: hepatoleitin A, scatter factor
C/Species: Homo sapiens (man)
C/Date: 17-Aug-1992 #sequence revision 17-Aug-1992 #ext change 09-Jul-2004
C/Accession: JH0579; J00333; A41140; B36677; A3512; A39006; PH0114; A37796; S06
R/Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A/Title: Organization of the human hepatocyte growth factor-encoding gene.
A/Reference number: JH0579; MUID:91340155; PMID:1831432
A/Accession: JH0579
A/Molecule type: DNA
A/Residues: 1-728 <SEK>
A/Cross-references: UNIPROT:P14210; DDBJ:D90318
A/Note: The authors translated the codon GAA for residue 662 as Gly
R/Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
Submitted to JIPID, March 1991
A/Description: Organization of the human hepatocyte growth factor-encoding gene.
A/Reference number: J00333
A/Accession: J00333
A/Molecule type: DNA
A/Residues: 1-481, RT, 484-728 <SE2>
R/Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A/Title: Evidence for the identity of human scatter factor and human hepatocyte growth f
A/Reference number: A41140; MUID:9134393; PMID:1831266
A/Accession: A41140
A/Molecule type: mRNA
A/Residues: 1-728 <WE1>
A/Cross-references: GB:M72339; NID:9337935; PIDN:AAA64239.1; PID:9337936
R/Seki, T.; Ihara, I.; Sugimura, A.; Shimomishi, M.; Nishizawa, T.; Asami, O.; Hagiya, i
Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A/Title: Isolation and expression of cDNA for different forms of hepatocyte growth fact
A/Reference number: A36677; MUID:91025062; PMID:2145836
A/Accession: B36677
A/Molecule type: mRNA
A/Residues: 1-728 <SE3>
A/Cross-references: GB:M60718; NID:9184031; PIDN:AAA52648.1; PID:9184032
A/Accession: A36677
A/Molecule type: mRNA
A/Residues: 1-161, 167-728 <SE4>
A/Cross-references: EMBL:X16323
A/Experimental source: leukocyte
R/Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakay
Biochem. Res. Commun. 163, 967-973, 1989
A/Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth fac
A/Reference number: A3512; MUID:89392017; PMID:2528952
A/Accession: A3512
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-728 <M1Y>

A:Cross-references: GB:M29145; NID:9184041; PIDN:AAA52650.1; PID:9306846
R:Rubin, J.S.; Chan, A.M.L.; Botcharo, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hid
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A:Title: A broad-spectrum human lung fibroblast-derived microgen is a variant of hepatoc
A:Reference number: A39006; MUID:9110540; PMID:1824873
A:Accession: A39006
A:Molecule type: mRNA
A:Residues: 1-161,167-728 <RUB>
A:Cross-references: GB:M55379
A:Experimental source: embryonic lung
R:Yoshiyama, Y.; Asakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama,
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A:Title: Identification of the N-terminal residue of the heavy chain of both native and
A:Reference number: PH0114; MUID:91207365; PMID:1826837
A:Accession: PH0114
A:Molecule type: protein
A:Residues: 32-43,53-58 <YOS>
A:Experimental source: plasma
R:Weidner, K.M.; Behrens, J.; Vandekekerckhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi
A:Reference number: A37796; MUID:91035621; PMID:2146276
A:Accession: A37796
A:Molecule type: protein
A:Residues: 86-91;329-344,356-363,'XX',366-370;425-434;442-447,'X',449-450;543-546,'X',5
R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimomishi, M.; Sugimura, A.; Tash
Nature 342, 440-443, 1989
A:Title: Molecular cloning and expression of human hepatocyte growth factor.
A:Reference number: S06794; MUID:90066676; PMID:2531289
A:Accession: S06794
A:Molecule type: mRNA
A:Residues: 1-31,'HK',34-77,'N',79-292,'V',294-299,'M',301-316,'A',318-335,'K',337-386,'
A:Cross-references: EMBL:X16323; NID:932081; PIDN:CAA34387.1; PID:932082
A:Experimental source: liver
A:Note: the authors translated the codon CAG for residue 727 as Glu
A:Note: part of this sequence, including the amino end of both the alpha and beta chains
R:Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
A:Reference number: I59214; MUID:93087571; PMID:1280830
A:Accession: I59214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-288,'ET' <HAR>
A:Cross-references: GB:I02931; NID:9184033; PIDN:AAA52649.1; PID:9184034
R:Miyaawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
Eur. J. Biochem. 197, 15-22, 1991
A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor g
A:Reference number: S15443; MUID:9120041; PMID:1826653
A:Accession: S15443
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288,'ET' <MY2>
A:Cross-references: EMBL:X57574; NID:932083; PIDN:CAA40802.1; PID:932084
R:Shima, N.; Nagaoka, M.; Ogaki, F.; Tsuda, E.; Muraikami, A.; Higashio, K.
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
A:Title: Tumor cytoxic factor/hepatocyte growth factor from human fibroblasts: cloning
A:Reference number: I52253; MUID:92062058; PMID:1835583
A:Accession: I52253
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 161-166 <SHI>
A:Cross-references: GB:S62561; NID:9237996; PIDN:AAB20169.1; PID:9237997
A:Accession: S62561
A:Molecule type: mRNA
A:Residues: 1-728 <SAS>
A:Cross-references: UNIPROT:Q00848; GB:D10212; NID:9220435; PIDN:BAAD1064.1; PID:9220436
A:Experimental source: fibroblast, COS-1 cell
A:Note: submitted to JIPID, May 1993
A:Accession: PC2064
A:Molecule type: protein
A:Residues: 496-504 <SA2>
R:Rosen, E.M.; Meirinsky, L.; Setzer, E.; Vinter, D.W.; Goldberg, I.D.
Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990

C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyr
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-494,495-728/Product: hepatocyte growth factor #status experimental <WAT>
F:332-494/Domain: alpha chain #status experimental <ACH>
F:128-206/Domain: kringle homology <KR1>
F:211-288/Domain: kringle homology <KR2>
F:305-383/Domain: kringle homology <KR3>
F:391-469/Domain: kringle homology <KR4>
F:495-728/Domain: beta chain #status experimental <BCH>
F:495-716/Domain: trypsin homology <TRY>
F:332/Modified site: pyroglutamate carboxylate acid (Gln) (in mature form) #status experimen
F:294,402,566,653/Binding site: carbonyl (Asn) (covalent) #status predicted
F:487-604/Disulfide bonds: #status predicted
Query Match 99.4%; Score 2563.5; DB 1; Length 728;
Best Local Similarity 98.7%; Pred. No. 1,66-175;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
QY 1 ERKRRNTHEFKSAKTTLLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCAFPYDK 60
DB 32 QKRRNTHEFKSAKTTLLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCAFPYDK 91
QY 61 ARKQCLMPFNSMSGYKKEGHEFDLYENKDYRNCIGKRGYKTVSTKSGICOP 120
DB 92 ARKQCLMPFNSMSGYKKEGHEFDLYENKDYRNCIGKRGYKTVSTKSGICOP 151
QY 121 WSMNIPHEH-----SYRGKDLQENYCRNPRGEGGPMCFSTNPEVREVCIDPOCSEYEC 175
DB 152 WSMNIPHEHSTLPSPSYRGKDLQENYCRNPRGEGGPMCFSTNPEVREVCIDPOCSEYEC 211
QY 176 MTCNGESYRGIMDHTESGKICQRMWDHQTPIHAKFLPERYPDKGFDNVCNRPDQCPRPWC 235
DB 212 MTCNGESYRGIMDHTESGKICQRMWDHQTPIHAKFLPERYPDKGFDNVCNRPDQCPRPWC 271
QY 226 YTLDPHRTMEYCAIKTCADNTMNDTDPVLETFTECIGOGSGYRGTVMTIMGIPQOMDS 295
DB 272 YTLDPHRTMEYCAIKTCADNTMNDTDPVLETFTECIGOGSGYRGTVMTIMGIPQOMDS 331
QY 296 QYPRHDMTPENPFCKDIRENYCRNPDGSESPPMCFITDPNIRVGYCSQIPNCDMSHGQDC 355
DB 332 QYPRHDMTPENPFCKDIRENYCRNPDGSESPPMCFITDPNIRVGYCSQIPNCDMSHGQDC 391
QY 356 YRGNGKNTMGLSOTRSGLTCSMDKNMEDLHRII FNEPDASKLNENYCRNPDDDAHGPW 415
DB 392 YRGNGKNTMGLSOTRSGLTCSMDKNMEDLHRII FNEPDASKLNENYCRNPDDDAHGPW 451
QY 416 CYTGNPLIPMDYCPISRCESGDTPTIV 442
DB 452 CYTGNPLIPMDYCPISRCESGDTPTIV 478
RESULT 2
A60185
hepatocyte growth factor precursor - mouse
N:Alternate names: hepatolein A; scatter factor
C:Species: Mus musculus (house mouse)
C:Dates: 03-Mar-1993 #sequence revision 26-May-1994 #text change 09-Jul-2004
A:Accession: J02117; PC2064; A60185; S43416; S45521; S17173; S10966; I48758; J00231
R:Saeki, M.; Nishio, M.; Saeki, T.; Enami, J
Biochem. Biophys. Res. Commun. 199, 772-779, 1994
A:Title: Identification of mouse mammary fibroblast-derived mammary growth factor as hep
A:Reference number: J02117; MUID:94183257; PMID:8135822
A:Accession: J02117
A:Molecule type: mRNA
A:Residues: 1-728 <SAS>
A:Cross-references: UNIPROT:Q00848; GB:D10212; NID:9220435; PIDN:BAAD1064.1; PID:9220436
A:Experimental source: fibroblast, COS-1 cell
A:Note: submitted to JIPID, May 1993
A:Accession: PC2064
A:Molecule type: protein
A:Residues: 496-504 <SA2>
R:Rosen, E.M.; Meirinsky, L.; Setzer, E.; Vinter, D.W.; Goldberg, I.D.
Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990

```

A>Title: Purified scatter factor stimulates epithelial and vascular endothelial cell mig
A:Reference number: A60185; MUID:50377927; PMID:2144630
A:Accession: A60185
A:Molecule type: protein
A:Residues: 'X',184-188,'KX',191-192,'X',194,'XX',197,357-364,'XX',367,375-377,'E',379,'
Riliu, Y.; Michalopoulos, G.K.; Zarnegar, R.
Biochim. Biophys. Acta 1216, 299-303, 1993
A>Title: Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth
A:Reference number: S43416; MUID:94060105; PMID:8241272
A:Accession: S43416
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-728 <LIU>
A:Cross-references: EMBL:X72307
Riliu, Y.
submitted to the EMBL Data Library, May 1993
A:Reference number: S45521
A:Accession: S45521
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-563,'H',565-728 <LI2>
A:Cross-references: EMBL:X72307
R.Coffier, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
Biochem. J. 278, 35-41, 1991
A>Title: Purification and characterization of biologically active scatter factor from ra
A:Reference number: S17173; MUID:51354223; PMID:1831975
A:Accession: S17173
A:Molecule type: protein
A:Residues: 496-517,'T',519 <COF>
R.Gherardi, E.; Stoker, M.
Nature 346, 228, 1990
A>Title: Hepatocytes and scatter factor.
A:Reference number: S10966; MUID:50326152; PMID:2142751
A:Accession: S10966
A>Status: preliminary
A:Molecule type: protein
A:Residues: 496-507,'X',509-512,'L',514-516,'X',518-519 <NAT>
R.Plaschke-Schlutter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.
J. Biol. Chem. 270, 830-835, 1995
A>Title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
A:Reference number: I48758; MUID:95122532; PMID:7822318
A:Accession: I48758
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-30 <RBS>
A:Cross-references: EMBL:X81630; NID:9673451; PIDN:CAAS7286.1; PID:9673452
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A>Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyz
F:1-32/Domain: signal sequence #status predicted <SIG>
F:56-495,496-728/Product: hepatocyte growth factor #status predicted <MAT>
F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:129-207/Domain: kringle homology <KR1>
F:212-289/Domain: kringle homology <KR2>
F:308-384/Domain: kringle homology <KR3>
F:392-470/Domain: kringle homology <KR4>
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F:496-719/Domain: trypsin homology <TRY>
F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:295,403,569,666/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:488-607/Disulfide bonds: #status predicted

Query Match 91.9%; Score 2370.5; DB 1; Length 728;
Best Local Similarity 89.9%; Pred. No. 9.9e-162;
Matches 402; Conservative 21; Mismatches 19; Indels 5; Gaps 1;

1 EKKRRNTTIEFKKSATTLTKIDPALKIKTKKVVNTADGCARCTRNGKLPFTCKAFVDPK 60
:::|||||
33 QKKRRNTTIEFKKSATTLTKEDPLKIKTKKVSADGCANRCJRNRPFTCKAFVDPK 92

```

QY	61	ARKOIMPEPMSMSGKYKKEGHEFDLYENKDYIRNCIIGKGRGKCTVSTTSGIKCOP	120
Db	93	SRKCCYWPENPSMSGKYKGFGEFDLYENKDYIRNCIIGKGGSKCTVSTTSGIKCOP	152
QY	121	WSSMI PHEH-----SYRGKDLQENYCRNPRGEEGPMCFTSNPEVRYEVCDI PCCSEVEC	175
Db	153	WNSMI PHEHSLPSSYRGKDLQENYCRNPRGEEGPMCFTSNPEVRYEVCDI PCCSEVEC	212
QY	176	MTNGESYRGLMDHTESGKICQRMWHOTPHRHKFLPERYPDKGFDDNYCRNPDQPRWC	235
Db	213	MTNGESYRGPMDHTESGKTCQRMWQOTPHRHKFLPERYPDKGFDDNYCRNPDGKPRWC	272
QY	236	YTLDPHTRMECAKTKCADNTMNDTVLPLETTELCIOQGGESYRGCTVNTIWNGLICQGRWS	295
Db	273	YTLDPTRPMECAKTKCHASAVNETDVPMETTELCIOQGGESYRGCTVNTIWNGLICQGRWS	332
QY	296	QYPHEHMDPTNPFCKDLRENYCRNPDGSESPMCFTTDPNIRVYCSQI PNCMDSHODC	355
Db	333	QYPHKHDTPTNPFCKDLRENYCRNPDGABSPMCFTTDPNIRVYCSQI PKCDVSSQDC	392
QY	356	YRGNGKNYMGLSQTRSGGLTCSMWDKNMEDLHRHIFWEPDASKLNTENYCRNPDDAHGPW	415
Db	393	YRGNGKNYMGLSKTRSGGLTCSMWDKNMEDLHRHIFWEPDASKLNTENYCRNPDDAHGPW	452
QY	416	CYTGNPLIPMDYCPISRCEGDTTPTIV	442
Db	453	CYTGNPLIPMDYCPISRCEGDTTPTIV	479

RESULT 3

hepatocyte growth factor precursor - rat
 N.Alternate names: hepatopietin A, scatter factor
 C.Species: Rattus norvegicus (Norway rat)
 C.Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 09-Jul-2004
 C.Accession: A35644, S13211
 R.Taahiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimoiishi, M.; Shimizu, S.; Nakamura,
 Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
 A.Title: Deduced primary structure of rat hepatocyte growth factor and expression of th
 A.Reference number: A35644, MUID:90222197, PMID:2139229
 A.Accession: A35644
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-728 <TRAS>
 A.Cross-references: UNIPROT:P17945; GB:D90102; GB:U32987; NID:9220766; PIDN:BA41433.1;
 A.Note: the authors translated the codon GAG for residue 70 as Gln, GAC for residue 417
 R.Okejima, A.; Miyazawa, K.; Kitamura, N.
 Eur. J. Biochem. 193, 375-381, 1990
 A.Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA du
 A.Reference number: S13211, MUID:91031482; PMID:2146117
 A.Accession: S13211
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-728 <OKA>
 A.Cross-references: EMBL:X54400; NID:956553; PIDN:CA48266.1; PID:94539554
 C.Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C.Function:
 A.Description: stimulates mitosis of hepatocytes and other cells
 A.Note: does not have proteinase activity
 C.Species: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
 C.Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyro
 F.1-32/Domain: signal sequence #status predicted <SIG>
 F.56-495/Product: hepatocyte growth factor #status predicted <MAT>
 F.56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F.129-207/Domain: kringle homology <KR1>
 F.212-289/Domain: kringle homology <KR2>
 F.306-384/Domain: kringle homology <KR3>
 F.393-470/Domain: kringle homology <KR4>
 F.496-728/Domain: hepatocyte growth factor beta chain #status predicted <BGH>
 F.496-719/Domain: trypsin homology <TRY>
 F.333/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte
 F.295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted
 F.488-607/Disulfide bonds: #status predicted

Query Match 91.6%; Score 2363.5; DB 1; Length 728;
Best Local Similarity 89.3%; Pred. No. 3.1e-161;
Matches 399; Conservative 23; Mismatches 20; Indels 5; Gaps 1;

```
QY 1 EKKRNTIHEFKKSAKTTLIKIDPALKITKKYKNTADOCANRCTRNGKLPTCKAFVFDK 60
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 33 OKRRNTLHEFKKSAKTTLIKIDPALKITKKYKNSADBCANRCKIRKNGPFTCKAFVFDK 92
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 61 AKOCIMFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 120
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 93 SKRCKWYFNFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 152
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 121 WSMIPIHEH-----SYRGDLOENYCRNPRGEGGPMCTSPNPEYRVEVCDIPQCEVEEC 175
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 153 WMSMIPHEHSFLPSSSYRGDLOENYCRNPRGEGGPMCTSPNPEYRVEVCDIPQCEVEEC 212
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 176 MTCNGESYRGDLMDHTESGKICORMDHQTPIRHKFLPERIPDYGFDDNYCRNPDGQPRPWC 235
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 213 MTCNGESYRGPMDHTESGKTCORMDQOTPIRHKFLPERYPDKGFDPNYCRNPDGKPRPWC 272
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 236 YTLDPHTREYCAIKTCADNTMNDTPVPLETTECIGOGEGYRGVTNTIWNGLPCORMDS 295
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 273 YTLDPHTREYCAIKTCASAVNETDVPETTECLIKGSGYRGVTNTIWNGLPCORMDS 332
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 296 QYRPHEDMTPENFCKDLRENYCRNPDGSGSPMCTTDPNIRVGYCSQIPNCMSHGQDC 355
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 333 QYRPHKDIRPENPKCKDLRENYCRNPDGSGSPMCTTDPNIRVGYCSQIPKCDVSGQDC 392
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 356 YGNGKNGYKNGNTSQTRSGITCSGMDKMDLHRHIFWEPDASKLBNYCRNPDDDAHGWP 415
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 393 YGNGKNGYKNGNLSKTRSGITCSGMDKMDLHRHIFWEPDASKLBNYCRNPDDDAHGWP 452
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 416 CYTGNDLIPMDYCPISRCGDDTPTIV 442
  ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 453 CYTGNDLIPMDYCPISRCGDDTPTIV 479
  ::::::::::: ::::::::::: ::::::::::: :::::::::::
```

RESULT 4

151283
hepatocyte growth factor precursor - clawed frog
N:Alternate names: heparinectin A; scatter factor
C:Species: Xenopus sp. (clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-May-2004
C:Accession: I51283
R:Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiohawa, K.
Mech. Dev. 49, 123-131, 1995
A:Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ear
A:Reference number: I51283; MUID:95267690; PMID:7748783
A:Accession: I51283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-710 <NAK>
A:Cross-references: GB:S77422; NID:9998932; PIDN:AA834354.1; PID:9998933
A>Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleoti
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A>Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1, kringle homolo
C:Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle
F:42-477/478-709/Product: hepatocyte growth factor #status predicted <Mat>
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <Ach>
F:115-193/Domain: kringle homology <KR1>
F:198-275/Domain: kringle homology <KR2>
F:289-367/Domain: kringle homology <KR3>
F:375-453/Domain: kringle homology <KR4>
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCh>
F:52-128,281,322,379,550,637,666/Binding site: carbohydrate (Aan) (covalent) #status pre
F:470-586/Disulfide bonds: #status predicted

Query Match 71.0%; Score 1830; DB 1; Length 710;

Best Local Similarity 68.4%; Pred. No. 3.8e-123;
Matches 301; Conservative 63; Mismatches 68; Indels 8; Gaps 2;

```
QY 3 KRRNTIHEFKKSAKTTLIKIDPALKITKKYKNTADOCANRCTRNGKLPTCKAFVFDK 62
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 21 KRRNAPFDYKKTAEITLRLKALBKVKTMNTTENCAKCRKNGKLPTCKAFVFDKNI 80
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 63 KOCIMFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOPWS 122
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 81 KRCHWFSFNTWSAGIKDYDISFDLYEKDYIRDCIHGKSGYNGRNTVTRKGLACOPWN 140
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 123 SMIPHEH-----YRGDLOENYCRNPRGEGGPMCTSPNPEYRVEVCDIPQCEVEECMT 177
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 141 SMIPHEHFLPSTYRGDLOENYCRNPRGEGGPMCTSPNPEYRVEVCDIPQCEVEECMT 200
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 178 CNGESYRGDLMDHTESGKICORMDHQTPIRHKFLPERYPDKGFDPNYCRNPDGQPRPWC 237
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 201 CNGEYRGPMDYTESGKICORMDLORPHKHFREBRIPNKGLNDNYCRNPDGKSPMCTT 260
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 238 LDPHTREYCAIKTCADNTMNDTPVPLETTECIGOGEGYRGVTNTIWNGLPCORMDSQY 297
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 261 LDPHTREYCAIKTCVHSHIVNNTDI--TKDCMKGOGEGYRGVSTTYNGIQCORMDSQF 317
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 298 PHEHDMTEBNKCKDLRENYCRNPDGSGSPMCTTDPNIRVGYCSQIPNCMSHGQDCYR 357
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 318 PHLNHFPEBNYKCKDLSENYCRNPDGSGSPMCTTDPNIRVGYCSQIPKCDVSGQDCY 377
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 358 GNGKNGYKNGNLSQTRSGITCSGMDKMDLHRHIFWEPDASKLBNYCRNPDDDAHGWP 417
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 378 GNGSYTKETLRTIRRLPCSMWEKRLQDLKHNTRNEPVSLQYKNCNPNDDAHGWPXY 437
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 418 TGNPLIPMDYCPISRCGDDT 437
  ::::::::::: ::::::::::: :::::::::::
Db 438 TDDPFPMDYCPISRCGDDT 457
  ::::::::::: ::::::::::: :::::::::::
```

RESULT 5

151285
hepatocyte growth factor/scatter factor - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: I51285
R:Steitl, A.; Stern, C.D.; They, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Cherard
Development 121, 813-824, 1995
A:Title: A role for HGF/SF in neural induction and its expression in Hensen's node durin
A:Reference number: I51285; MUID:95237013; PMID:7720585
A:Accession: I51285
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-411 <STR>
A:Cross-references: UNIPROT:Q90978; GB:S77480; NID:9998675; PID:9998676
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1, kringle homolo
F:124-197/Domain: kringle homology <KR1>
F:202-279/Domain: kringle homology <KR2>
F:296-374/Domain: kringle homology <KR3>

Query Match 67.9%; Score 1750; DB 2; Length 411;
Best Local Similarity 75.9%; Pred. No. 1.1e-117;
Matches 289; Conservative 47; Mismatches 45; Indels 0; Gaps 0;

```
QY 3 KRRNTIHEFKKSAKTTLIKIDPALKITKKYKNTADOCANRCTRNGKLPTCKAFVFDK 62
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 30 KRRNPLHDYKKTGELMLIKVNTLEVKTKLNTTEQCCKRSRNGKLSFTCKAFVADVT 89
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 63 KOCIMFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOPWS 122
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 90 KRCHWLSNLSJLNGYRKQDAFPLFEKKDVRNRCIIIGKGEYKGTISTISGICQAMN 149
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 123 SMIPHEHSGDLOENYCRNPRGEGGPMCTSPNPEYRVEVCDIPQCEVEECMT 182
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 150 SMIPHEHSGDLOENYCRNPRGEGGPMCTSPNPEYRVEVCDIPQCEVEECMT 209
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 183 YRGDLMDHTESGKICORMDHQTPIRHKFLPERYPDKGFDPNYCRNPDGQPRPWCCTLDPT 242
  ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
```

Db 210 YGPMPTHTSGKCEQCRMDLQRPKHKFRBERYPDKFFDNNYCNRPDGKLRPMWCYTLDPT 269
QY 243 RMEYCAIKTCADNTMDTVPLETTECTICGGGEGYGTWNTWNGIPCORMDSQYHEHD 302
Db 270 PMEFCAIKTCIDVGLNSTEVAETTTTCICGGGEGYGTWNTWNGIPCORMDSQYHEHD 329
QY 303 MTPENFKCDLRENNYCRNPDGSESPMCFTTDPNIRVYGSQIPNCMDSHQDCYRNGKN 362
Db 330 ITPENFKCDLRENNYCRNPDGSESPMCFTTDPNIRVYGSQIPNCMDSHQDCYRNGKN 389
QY 363 YMGNSQTRSGLTCSMWDXM 383
Db 390 YMGNSKTRFGLTCSMTDXM 410

RESULT 6
A47136
macrophage-stimulating protein 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence, revision 14-Nov-1997 #text_change 09-Jul-2004
C:Accession: A40331; A40331; A47136; A61395
R:Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A:Title: Characterization of the DNFI52 locus on human chromosome 3: identification of
A:Reference number: A40331; MUID:92002016; PMID:1655021
A:Accession: A40331
A:Molecule type: DNA
A:Residues: 1-711 <HA1>
A:Cross-references: UNIPROT:P26927; GB:M74179
A:Accession: B40331
A:Molecule type: mRNA
A:Residues: 1-711 <HA2>
A:Cross-references: GB:M74178; NID:g183976; PIDN:AA50165.1; PID:g183977
R:Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A:Title: Cloning, sequencing, and expression of human macrophage stimulating protein (MS
A:Reference number: A47136; MUID:93340141; PMID:8393443
A:Accession: A47136
A:Molecule type: mRNA
A:Residues: 1-12, 'C', '14-622', 'F', '624-711' <YOS>
A:Cross-references: GB:U1924; NID:g398037; PIDN:AA59872.1; PID:g398038
A:Note: authors translated the codon TTT for residue 623 as Leu; parts of this sequence
R:Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 173, 1227-1234, 1991
A:Title: Macrophage stimulating protein: purification, partial amino acid sequence, and
A:Reference number: A61395; MUID:91217635; PMID:162741
A:Accession: A61395
A:Molecule type: protein
A:Residues: 230-247,288-291, 'E', '293-295', 'X', '297-301', 'X', '303', 'E', '305', 'EX', '308-310,326-331
A:Experimental source: plasma
C:Genetics:
A:Gene: GDB:MST1, D3P15S2, DNFI5S2, HGFL
A:Cross-references: GDB:128833; OMIM:142408
A:Map position: 3p21-3p21.3
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
C:Keywords: duplication; glycoprotein; growth factor; kringle; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-483,484-711/Product: macrophage-stimulating protein 1 #status predicted <MA>
F:19-483/Domain: alpha chain #status predicted <ACH>
F:110-186/Domain: kringle homology <KR1>
F:191-268/Domain: kringle homology <KR2>
F:283-361/Domain: kringle homology <KR3>
F:370-448/Domain: kringle homology <KR4>
F:484-711/Domain: beta chain #status predicted <BCH>
F:484-704/Domain: trypsin homology <TRY>
F:56-78,66-110-166,131-169,157-181,191-268,212-251,240-263,283-361,304-343,332-355,37
F:72,296,315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.4%; Score 1171; DB 1; Length 711;
Best Local Similarity 47.3%; Pred. No. 4,4e-76;
Matches 196; Conservative 60; Mismatches 148; Indels 10; Gaps 4;

QY 33 VNTADQANRCRTRNKGLPFTCKAFVDFDKARQCLWPFNSMSSGVKKEFGHEFDLYENKD 92
Db 50 VADAEBCAGRC-----GPLMDCAFFHYNNVSSHCCQLLPWTOHSHPTLRASGRCDLPQKXD 105
QY 93 YTRNCTIGKRGYKGTSTITKSGIKCQPMSSMI PHEHSR---GKDLOENYCRNPRGEG 149
Db 106 YVRTCIMNNGVYRGMTATVGGILPCQAMSHFPPNDHKYPTLRNGLEENFCRNPDGDFQ 165
QY 150 GPMCFSTNPEVYEVCDIPQCEVECMTCNGESYRGLMDHTESGKICQEMDQTPRHKE 209
Db 166 GPMCTTDDPAVAFQSGCIGSCREAAACWNGEYTGANDRTSGRECCQWMDQHPQHF 225
QY 210 LPERYPDKGFDDNNYCRNPDGQPRPWCYTLDPTHTREYCAIKTCADNTMDTVPLETTEC 269
Db 226 EPKFLDQGLDDNNYCRNPDGSESPMCYTLDPTIEREFCOLPRCGSBAQROGA--TTVSC 283
QY 270 ICGGEGYGTWNTWNGIPCORMDSQYHEHDTPENFKCDLRENNYCRNPDGSESPMC 329
Db 284 FPGKGGYGTWNTWNGIPCORMDAQIPQHRFTPEKYACDLRENNYCRNPDGSESPMC 343
QY 330 FTTDPNIRVYGSQIPNC-DMSHGDCYRGNKYNWNLSTRTSGLTCSMDKXMDLHR 388
Db 344 FTLRPEMRAPCYQIRCTDVRPOCTYGAAGEQIRGTYSKTRKGVQCORMSAETPRHKE 403
QY 389 HIFWEPDASKLENYCRNPDGDAHGPWCYTGPNLIPWDYCPISRCBGTPTIV 442
Db 404 FFTSEPHAQLENNYCRNPDGSHGPMWCYTMDPRPFVDCALRCADDQPSIL 457

RESULT 7
JC5061
macrophage-stimulating protein 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Jan-1997 #sequence, revision 31-Jan-1997 #text_change 09-Jul-2004
C:Accession: JC5061
R:Oshimuro, K.; Iwano, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, T.
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A:Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
A:Reference number: JC5061; MUID:97011126; PMID:8858136
A:Accession: JC5061
A:Molecule type: mRNA
A:Residues: 1-716 <ONS>
A:Cross-references: P0521; EMBL:X95096; NID:g1669718; PIDN:CAA64473.1; PID:g166
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolo
C:Keywords: duplication; glycoprotein; growth factor; kringle
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MA>
F:32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F:110-186/Domain: kringle homology <KR1>
F:191-268/Domain: kringle homology <KR2>
F:292-370/Domain: kringle homology <KR3>
F:379-457/Domain: kringle homology <KR4>
F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F:489-709/Domain: trypsin homology <TRY>
F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.0%; Score 1133.5; DB 1; Length 716;
Best Local Similarity 46.1%; Pred. No. 2.1e-73;
Matches 195; Conservative 68; Mismatches 141; Indels 19; Gaps 6;

QY 33 VNTADQANRCRTRNKGLPFTCKAFVDFDKARQCLWPFNSMSSGVKKEFGHEFDLYENKD 92
Db 50 VADAEBCAGRC-----GPLMDCAFFHYNNVSSHCCQLLPWTOHSHPTLRASGRCDLPQKXD 105
QY 93 YTRNCTIGKRGYKGTSTITKSGIKCQPMSSMI PHEHSR---GKDLOENYCRNPRGEG 149
Db 106 YVRTCIMNNGVYRGMTATVGGILPCQAMSHFPPNDHKYPTLRNGLEENFCRNPDGDFQ 165
QY 150 GPMCFSTNPEVYEVCDIPQCEVECMTCNGESYRGLMDHTESGKICQEMDQTPRHKE 209
Db 166 GPMCTTDDPAVAFQSGCIGSCREAAACWNGEYTGANDRTSGRECCQWMDQHPHSHF 225

Qy	210	LBREYRDPGRPDNDYCRNPDQGRPRWCYTLLDHTNMEYALITCADN-----IMNDITDV	262
Db	226	HPEKFRDALKDNCRCNPDASERFWCYTTDNRVEREPFDLSGCGNPLRPPTTKSGSQORN	285
Qy	263	PLETTTECTIOQGEQEGYRGTVNTIWNIGPCORWDSQYRHEHDMTPRNFCKDLRENYCNP	322
Db	286	KVKASNCPRKGEDYRGTTTNTTSAGVPCQRMWDAQPHQHRVLPREKACCKDLRENFCCNP	345
Qy	323	GSESPWCTTTPNIRIVGCSQIPLNC-DMSHODCYRGNKRYMNLSTREGLTCSMWDK	381
Db	346	GSEAPWCTSTRPGRLFAVCYQIPRCTEEVVEEGCHSGEYRKSVSSTRGAVCCQHN--	403
Qy	382	NMEDLHRHIFWEPRD--SKLNENYCRNPDQDAHGWCYTGNPLIPMDYCPISRCGDTTP	439
Db	404	SSEIRPHKQFPRTSPAHNGLEANFCRNPDGSHRGWCYTLLDPTLLFDYCALKRCDDDQRP	463
Qy	440	TIV	442
Db	464	SIL	466

RESULT 8

macrophage-stimulating protein 1 precursor - mouse
N/Alternate names: hepatocyte growth factor-like protein
C/Species: Mus musculus (house mouse)
C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C/Accession: A40332; B40332
C/Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
Biochemistry 30, 9781-9791, 1991
A/Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor
A/Reference number: A40332; MUID:92002017; PMID:1832957
A/Accession: A40332
A/Molecule type: DNA
A/Residues: 1-1716 <DEG>
A/Cross-references: UNIPROT:P26928; GB:M4180; PID:G193833; PID:AAA50166.1; PID:G1938332
A/Accession: B40332
A/Molecule type: mRNA
A/Residues: 1-18,'P',20-716 <DEG2>
A/Cross-references: GB:M4181; NID:G193833; PID:AAA50167.1; PID:G193834
C/Genetics:
A/Introns: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2; 458/1; 47
C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C/Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homoloid
C/Keywords: duplication; glycoprotein; growth factor; kringle
F/1-11/Domains: signal sequence #status predicted <SIG>
F/19-468,469-716/Product: macrophage-stimulating protein 1 #status experimental <MGT>
F/19-469/Domains: alpha chain #status experimental <ACH>
F/110-186/Domains: kringle homology <KR1>
F/191-268/Domains: kringle homology <KR2>
F/292-370/Domains: kringle homology <KR3>
F/379-457/Domains: kringle homology <KR4>
F/484-711/Domains: beta chain #status experimental <BCH>
F/489-709/Domains: trypsin homology <TRY>
F/72-173,305,620/Binding site: carboxydrase (Asn) (covalent) #status predicted
F/72-173,305,620/Binding site: carboxydrase (Asn) (covalent) #status predicted

Query	March	43.2%	Score	1114.5;	DB	1;	Length	716;	
Best Local Similarity		45.9%*	Pred.	No. 4.8e-72;					
Matches	194;	Conservative	67;	Mismatches	145;	Indels	19;	Gaps	6;
Dy	33	VNTADQCNRCTRNKGLPFTCKAFVDFDARAKQCLFPPNSMSSGVYKEFGHEFDLYEKD	92						
		: : : : :	:	:	:	:	:	:	:
Dd	50	VADAECCARRC----	GPILDCRAFHYNMSSHGCCQLLPPTQSHLTQLYHSSLCHLFQAKD	105					
		: : : : :	:	:	:	:	:	:	:
Dy	93	YIRNCIIKGSRYSKGTVAITSRGIKCPWSSSMIPHEHSYR--GKDLOENYCRRNPGEEG	149						
		: : : : :	:	:	:	:	:	:	:
Dd	106	YVRTCIMONGVSIRSTVARTAGGLPCQAMSRFRPDHDKYTTPPKNGLEENCRNPDGDPR	165						
		: : : : :	:	:	:	:	:	:	:
Dy	150	GPWCETSPPEAYREYVCDIPOCSVEEVCMTGNESITRYGLMDHTESGITCORMDHQTPRHKKF	209						
		: : : : :	:	:	:	:	:	:	:
Dd	166	GPWCATTNRSYRFOSCGIKTCREAIVCYLNGEDVATGEADVATESGECRCRMLOJPHSHPF	225						
		: : : : :	:	:	:	:	:	:	:

QY	210	..PEEYPRKGFDDNNCRNPDDGQPRMWCYTLDPHPRMWECAIKTCAADN-----TMDTDV	262
Db	226	QPEKFLDKDLKDNCRNPDDGSERPMCYTTPNNVEREFCDLPSGGPNLPPIVVKSSKSRN	285
QY	263	PLETTTECIQGGEGEGYRGTVNTIMNGIIPQQRWDSQPPHEHDMTPENFKCKDLRENYCRNP	332
Db	286	KGKLIANFCRFGGEGEDYRGTTTTSAGVPCQRWDQSPHQHIFVEBKVACKDLRENFCRNP	345
QY	323	GSESPPMCFETDPNIRVGYGSGOIPWC--DMSHGDCYCRNGKNGVNGNLSQTSGLTCSMWDX	381
Db	346	GSEAPMCFETSPGIRRMFCQIPIRCTEELVPEGCTHSGEGYQGSVKTKKGQQCQHW--	403
QY	382	NMEDLHRHIFMEPDASK--LNENYCRNPDDDAHGPWCYTGNPLIPWDYCPISRCCEGDTTP	439
Db	404	SSEIRPHKQFIPTSPAPAGLEANFCRNPDDGSHGPWCYTLDPLIDILFDYCALQRCDDQPR	463
QY	440	TIV	442
Db	464	STL	466

RESULT 9

plasmin (EC 3.3.21.7) precursor - pig (fragment)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Sep-1990 #sequence revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: S03733; S03737; A25834
R:Schaller, J.; Marti, T.; Roessellet, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the cat
A:Reference number: S03733
A:Accession: S03733
A:Molecule type: protein
A:Residues: 1-560 <SCH>
A:Cross-references: UNIPROT:P06867
R:Brunholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03735, MUID:81212097; PMID:7228497
A:Accession: S03737
A:Molecule type: protein
A:Residues: 1-57 <BRU>
R:Marti, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A:Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.
A:Reference number: A25834; MUID:85203907; PMID:3846533
A:Accession: A25834
A:Molecule type: protein
A:Residues: 450-790 <MAR>
A:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
as the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine prote
F:1-790/Product: plasminogen #status predicted <PRO>
F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>
F:1-77/Domain: activation peptide #status predicted <APT>
F:78-560/Product: plasmin chain A #status predicted <ACH>
F:84-162/Domain: kringle homology <KR1>
F:166-243/Domain: kringle homology <KR2>
F:258-333/Domain: kringle homology <KR3>
F:358-435/Domain: kringle homology <KR4>
F:450-790/Product: miniplasminogen #status experimental <MIN>
F:461-540/Domain: kringle homology <KR5>
F:561-790/Product: plasmin chain B #status experimental <BCH>
F:561-783/Domain: trypsin homology <TRY>
F:30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,305
bonds: #status predicted
F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 40.9%; Score 1054.5; DB 1; Length 790;
Best Local Similarity 44.6%; Pred. No. 1e-67;
Matches 195; Conservative 64; Mismatches 133; Indels 45; Gaps 14;

25 ALKIKTKVNTADOCANRCCTRNKGLPTCKAFVVDKARKOCOLMPFNSMSGVKKERGHE 84
16 SLRKQVAASVEBCAKCAEETN--FICRAFYHSDQOCVMAENSKTSPIRM--RD 71
85 FDLYENKDYIRNCTIGKRSYKGVTSITKSGIKQCPWSSMI PHEHSYRGK----DLQEN 139
72 VVLFKRIYLSCEKGTGKNGYRGTTSKTSGVLCQKMSVSPH PKSPKPLAGLEEN 131
140 YCANNPGEEGGPGCFTSNPEVRYEVCIPQCESEVCHTCNGESYRGLMDHTESGKTCQRW 199
132 YCANNPDDEKGPWCYTTPDPTRFDYCIPEC--EDECHCSGEHHEGKISTKMSGIEQSW 190
200 DHQTPHRHKFLPERYPDKGFDNCRNPDGPRWCYTLDPHTRWEXCAIKTCADNTMD 259
191 GSQSPHAGVLPSPKFPKNTKNNYCNRPDEPRPWCFTTDPNKRWEFCDI-TPPP 248
260 TDVPLETTECTIOGQEGYRGTVNTIWMGICPQRMDSQYPHEHDMTPENFKCDLRENYCR 319
249 TSGP--TYQCLKRGEGYRGTIVSVTASGHTCQRMASQSPKHNRTPEPFCKLSENYCR 306
320 NPDGSSPWCFTTDPNIRVGYCSQIPNCDS-----H-----GQDCYRGNKN 362
307 NPDGETAPWCYTTDSEVRMDYC--KIPSCGSSITSTEHLADPVPDEQTPVAQDCYRNGES 365
363 YMCNLQTSRGLCSMWDMKMEDLHRIHFE-----PDASKLENYCRNPDADAHGPMWC 416
366 YRGTSTTTTGKRCQSVSMTP--HRH--EKTGPNFPNAG-LTMNCRNPDAD-KSPWC 418
417 YTGNPILPMDYCEISRC 433
419 YTTDPVRWEXYCNLKKC 435

Db

RESULT 10
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N:Contains: angiotatin; plasminogen
C:Species: Mus musculus (house mouse)
C>Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C/Accession: A38514; S48202; S48203
R/Begun, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A>Title: Characterization of the cDNA coding for mouse plasminogen and localization of
A/Reference number: A38514; MUID:91184812; PMID:2081600
A/Accession: A38514
A/Molecule type: mRNA
A/Residues: 1-812 <DEG>
A/Cross-references: UNIPROT:P20918; GB:J04766; NID:g200402; PIND:AAA50168.1; PID:g200403
R/Lijjen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A>Title: Characterization of the murine plasma fibrinolytic system.
A/Reference number: S48202; MUID:95010076; PMID:7523120
A/Accession: S48202
A/Molecule type: protein
A/Residues: 20-25
A/Accession: S48203
A/Molecule type: protein
A/Residues: 22-27
C/Comment: Plasminogen is synthesized by the kidney and is present in plasma and many of
C/Comment: plasminogen is converted into plasmin by plasminogen activators, both plasmin
C/Comment: immediately after dissociation from the clot. In the presence of the inhibitor, the activa
C/Comment: e inhibitor, the activation involves also removal of the activation peptide.
C/Comment: Stremelysin 1 (see PIR:KCMSI) acts on plasminogen to produce angiotatin. To
C/Function: eful in treating solid tumors.
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
A/Pathway: fibrinolysis
C/Superfamily: plasmin; kringie homology; plasminogen-related protein precursor homology

C/Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyd
F/1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-812/Product: plasminogen #status predicted <PRO>
F/20-96/Domain: activation peptide #status predicted <APT>
F/79-466/Product: angiotatin #status predicted <AST>
F/97-581,582-812/Product: plasmin #status predicted <MAT>
F/97-581/Domain: chain A #status predicted <ACH>
F/103-181/Domain: kringie homology <KR1>
F/185-262/Domain: kringie homology <KR2>
F/275-352/Domain: kringie homology <KR3>
F/377-454/Domain: kringie homology <KR4>
F/461-560/Domain: kringie homology <KR5>
F/582-812/Domain: chain B #status predicted <BCH>
F/582-805/Domain: trypsin homology <TRY>
F/49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,3
bonds: #status predicted
F/78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F/136,308/Binding site: carboxylate (Asn) (covalent) #status predicted
F/466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F/581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F/624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 40.3%; Score 1039.5; DB 1; Length 812;
Best Local Similarity 43.1%; Pred. No. 1.2e-66;
Matches 190; Conservative 64; Mismatches 144; Indels 43; Gaps 13;

30 TKKVNTA---DQANCTRNKGLPTCKAFVFPKAKKOCIMFPNMSGVKKERGFEPD 86
37 TKQLAAGVSDCLACEBE--TDVCRSFQYHSKEQOCVMAENSKTSSIIIM--RDVI 92
87 LYENKDYIRNCTIGKRSYKGVTSITKSGIKQCPWSSMI PHEHSYR-----GDLQENYC 141
93 LPEKRYVLSCEKGTGKNGYRGTTSKTSGVLCQKMSVSPH PKSPKPLAGLEENYC 152
142 RNPGEGBGPGCFTSNPEVRYEVCIPQCESEVCHTCNGESYRGLMDHTESGKTCQRWD 201
153 RNPNDGPGWCYTTDPPDRYDCNIPCEBE--ECWYCSGEGYKISTKMSGIDCAWDS 211
202 QPFRHKFLPERYPDKGFDNCRNPDGPRWCYTLDPHTRWEXCAIKTCADNTMD 261
212 QSPHAGVLPSPKFPKNTKNNYCNRPDEPRPWCFTTDPNKRWEFCDI-TPPP 267
262 VPLETTECTIOGQEGYRGTVNTIWMGICPQRMDSQYPHEHDMTPENFKCDLRENYCRN 321
268 PEPPTVQCLKRGEGYRGTIVSVTASGHTCQRMASQSPKHNRTPEPFCKLSENYCRN 327
322 DGSESPWCFTTDPNIRVGYCSQIPNCDSHG-----ODCYRGNKNYM 364
328 DGETAPWCYTTDSEVRMDYC--KIPSCSSASPDQSSVPEBQTPVQECYSDQSYR 386
365 GNLSQTSRGLCSMWDMKMEDLHRIHFE-----PDASKLENYCRNPDADAHGPMWC 421
387 GTSSTTTTGKRCQSVSMTP--HRHKTPEPFCKLSENYCRN 442
422 LIPMDYCEISRCGDTPTIV 442
443 SVRWEXYCNLKKC-SETGGSV 462

Db

RESULT 11
146260
plasmin (EC 3.4.21.7) precursor - western European hedgehog
C/Species: Erinaceus europaeus (western European hedgehog)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C/Accession: I46260
R/Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fon
J. Biol. Chem. 270, 24004-24009, 1995
A>Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipopro
A/Reference number: I46259; MUID:96025778; PMID:7592597
A/Accession: I46260
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA

A:Residues: 1-810 <LAW>
 A:Cross-references: UNIPROT:Q29485; EMBL:U33171; NID:g1046360; PID:g1046361
 C:Superfamily: plasmin: kringle homology; plasminogen-related protein precursor homology
 C:Keywords: hydrolase; serine proteinase
 F:103-181/Domain: kringle homology <KR1>
 F:185-262/Domain: kringle homology <KR2>
 F:275-352/Domain: kringle homology <KR3>
 F:379-456/Domain: kringle homology <KR4>
 F:482-561/Domain: kringle homology <KR5>
 F:582-803/Domain: trypsin homology <TRY>

Query Match 40.2%; Score 1036; DB 2; Length 810;
 Best Local Similarity 43.8%; Pred. No. 2,2e-66;

Matches 192; Conservative 64; Mismatches 134; Indels 48; Gaps 14;

```

QY 30 TKK--VNTADQANRCRNKGLPFTCKAFVFDKARKQCLPFPMSSSGYKKEGFHFD 86
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 TKKQLAGSTIECAKACEE--EFTCRSFQYHSKEQOCVMAENKRSIYFRM--RDV 92
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 LYENKDYIRNCIIIGKGRSVITKSGIKQCPWSSMIPHEHSYR-----GKDLQENYC 141
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 LFEKMYLSECKVGNKGYRGTVSKTGTGLTCQKMSAETPHKPRSPDENPEGLDQNYC 152
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 142 RNPGEBCGPMCFSTNPVRYEVCDIPOCEVECTCNAGESYRGLMDHTESGKICQAWDS 201
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 RNPNDPGPMCYTMDPEVREYCEIIOC-EDECHNCSQNYVGKISRTMSGLECQPMDS 211
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 202 QTPRHKFLPERYPDKGFDNYCRNPDGQRPWCYTLDPHTRWEYCAIKTCADNTMNDT 261
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 QTPHGHGTPSKFPKLNKMYCRNPDGPRPWCFTMDENKWEYCDIPRCT-----TP 265
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 262 VPLE--TTECIQOGEGYRVTNTIWNGLPCQRMDSQYRHEHDMTPENFKCKDLRENYCR 319
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 PPSGPTQYCLKNGENRGDAVATVSGTCHGMSAQTPHTHNTPEPFCKNDENYCR 325
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 NPDGSESPWCTTDPNIRVYCSQIPNCDMSH-----GDCYRANGK 360
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 NPDGAPAPWCFTTNSVREYFC-KIPDCVSSASETHSDAPVTPPEQTPVVECYQNG 384
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 KNYMNLSTRSGLTCSMDKXMDLHRIIWE---PDASKLNENYCRPDDAIGPMY 417
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 QTYRGTSSTTTGKKCQPMWTSMP--HRHSKTPENYPDAD-LTMVYCRNPDG-KGPMCY 440
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 418 TGNPLIPMDYCPISRCEG 435
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 441 TTDPSEYRWFENLKKCSG 458
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

B30848

plasmin (EC 3.4.21.7) precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004

C:Accession: B32869; B30848

C:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.

J. Biol. Chem. 264, 5957-5965, 1989

A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.

A:Reference number: A32869; MUID:89174660; PMID:2925643

A:Accession: B32869

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-810 <TOM>

A:Cross-references: UNIPROT:P12545; GB:J04697; NID:g342272; PID:AAA6901.1; PID:g342272

C:Superfamily: Plasmin: kringle homology; plasminogen-related protein precursor homology

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-96/Domain: plasminogen-related protein precursor homology <PLPB>

F:103-181/Domain: kringle homology <KR1>

F:185-262/Domain: kringle homology <KR2>

F:275-352/Domain: kringle homology <KR3>

F:377-454/Domain: kringle homology <KR4>

F:481-560/Domain: kringle homology <KR5>

F:581-803/Domain: trypsin homology <TRY>
 F:49-73,53-61,103-181,124-154,152-176,185-262,188-316,206-245,234-257,275-352,296-335,332
 bonds: #status predicted
 F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 39.6%; Score 1021; DB 2; Length 810;
 Best Local Similarity 42.2%; Pred. No. 2.6e-65;

Matches 186; Conservative 72; Mismatches 127; Indels 56; Gaps 13;

```

QY 30 TKK--VNTADQANRCRNKGLPFTCKAFVFDKARKQCLPFPMSSSGYKKEGFHFD 86
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 TKKQLAGSTIECAKACEE--EFTCRSFQYHSKEQOCVMAENKRSIYFRM--RDV 92
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 LYENKDYIRNCIIIGKGRSVITKSGIKQCPWSSMIPHEHSYR-----GKDLQENYC 141
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 LFEKMYLSECKVGNKGYRGTVSKTGTGLTCQKMSRSPHPRFSPATHSSELENYC 152
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 142 RNPGEBCGPMCFSTNPVRYEVCDIPOCEVECTCNAGESYRGLMDHTESGKICQAWDS 201
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 RNPNDGQPMCYTMDPEEREDYCDIPBC-EDECHNCSGENYDKISRTMSGLECQAWDS 211
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 202 QTPRHKFLPERYPDKGFDNYCRNPDGQRPWCYTLDPHTRWEYCAIKTCADNTMNDT 261
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 QTPHGHGTPSKFPKLNKMYCRNPDGPRPWCFTMDENKWEYCDIPRCT-----TP 265
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 262 VPLE--TTECIQOGEGYRVTNTIWNGLPCQRMDSQYRHEHDMTPENFKCKDLRENYCR 319
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 PPSGPTQYCLKNGENRGDAVATVSGTCHGMSAQTPHTHNTPEPFCKNDENYCR 325
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 NPDGSESPWCTTDPNIRVYCSQIPNCDMSH-----GDCYRANGK 362
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 NPDGAPAPWCFTTNSVREYFC-KIPDCSSPVSTEPDTPAPPBLPVPVVECYHGDQS 384
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 363 YGMNLSTRSGLTCSMDKXMDLHRIIWEFDPASKLNENYCRPDDAIGPMY 414
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 YRGTSSTTTGKKCOSMSMTPHMHEKTPENF-----PNAAG-LTMVYCRNPDG-KG 435
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 415 WCYTGNPLIPMDYCPISRCEG 435
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 WCFTTDPSEYRWFENLKKCSG 456
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 13

PLB0

plasmin (EC 3.4.21.7) precursor - bovine

N:Alternate names: plasminogen

C:Species: Bos primigenius taurus (cattle)

C:Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004

C:Accession: S45046; A25835; I45961; S03736

R.Berglund, L.; Andersen, M.D.; Petersen, T.E.

submitted to the EMBL Data Library, May 1994

A:Description: Cloning and characterization of the bovine plasminogen cDNA.

A:Reference number: S45046

A:Accession: S45046

A:Molecule type: mRNA

A:Residues: 1-812 <BER>

A:Cross-references: UNIPROT:P06868; EMBL:X79402; NID:g494962; PID:CAA55939.1; PID:g494949

A:Experimental source: liver

A:Note: it is uncertain whether Met-1 or Met-8 is the initiator

R.Schallier, J.; Moser, P.W.; Dammegeer-Muller, G.A.K.; Rosselet, S.J.; Kamfer, U.; Rick

Eur. J. Biochem. 149, 267-278, 1985

A:Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plas

A:Reference number: A25835; MUID:85203906; PMID:3846532

A:Accession: A25835

A:Molecule type: protein

A:Residues: 27-334,'D',336-515,'H',517-554,'L',556-812 <SCH>

R.Malinowski, D.P.; Sadler, J.E.; Davie, E.W.

Biochemistry 23, 4243-4250, 1984

A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and

A:Accession: I45961

A:Status: translated from GB/EMBL/DBD

A:Molecule type: mRNA

A:Residues: 706-743, 'R', 745-812 <MAL>
A:Cross-references: GB:K02935; NID:9163551; PIDN:AAA0714.1; PID:g163552
R:Brumsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1991
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03736
A:Molecule type: protein
A:Residues: 27-83 <BRU>
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringe homology; plasminogen-related protein precursor homology
C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringe; plasma;
F:1-26/Domain: signal sequence #status predicted <SIG>
F:8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F:27-812/Product: plasminogen #status experimental <PRO>
F:27-103/Domain: activation peptide #status experimental <APR>
F:104-583, 584-812/Product: plasmin #status experimental <MAT>
F:104-583/Domain: plasmin chain A #status experimental <ACH>
F:110-188/Domain: kringe homology <KR1>
F:192-269/Domain: kringe homology <KR2>
F:282-359/Domain: kringe homology <KR3>
F:384-461/Domain: kringe homology <KR4>
F:485-564/Domain: kringe homology <KR5>
F:584-812/Domain: plasmin chain B #status experimental <BCH>
F:584-805/Domain: trypsin homology <TRY>
F:56-80, 60-68, 110-188, 131-171, 159-183, 192-269, 195-323, 213-252, 241-264, 282-359, 303-342, 33
Bonds: #status predicted
F:315/Binding site: carbonylate (Asn) (covalent) #status experimental
F:365/Binding site: carbonylate (Ser) (covalent) #status experimental
F:624, 667, 762/Active site: His, Asp, Ser #status predicted
Query Match 39.5%; Score 1019.5; DB 1; Length 812;
Best Local Similarity 42.0%; Pred. No. 3.3e-65;
Matches 183; Conservative 73; Mismatches 137; Indels 43; Gaps 12;
25 ALKIKKKVNTAQCANRCRNKGLPPTCKAFVDFDARAKOCMLFPNSMSSGKKEFGHE 84
42 SLRKKLAGRSVDCAAKC--EETPFVCAFAFYHSKEQCVMMAENSKATPVFRM--R 97
85 FDLYENKDYIRNCLIGKGRSYKTVASITSGIKCQPMSSMI PHEHSYRGK-----DLOEN 139
98 VILYERKIRILLBEKTONGQYRTGTATETSGVTCQMSATSPHVPKFPESPEKFLAGLEEN 157
140 YCRNPGEERGCPWCTSNPEVRYECDDIPQCSVEECMTGNGESYRGLMDHTESGKICQRY 199
158 YGNPNPNDENGPCWCTTDPDKRYDCDIEC--EDKCMHSGENYEKIAKTMSGRDCQAW 216
200 DHQTPHRRKFLPRRYDCKGDDNYCCRNPGQCPKWCYTLDPHRMRYCAIKTCADNTMND 259
217 DSQSPRAHGITISKFPNKKLKKYCNKPNQDEPRPWCFTTDPQKWEFCDPRT----- 270
260 TVDPLE--TTECTIOGEGEYRGTVNTIWNIGIPCCQWDSQYPRHMDTPENFKCKDLRENY 317
271 TTPPSSGPKYQCGLKGYKAVYVTSBHTCQKMSSEQPHHNTPTENFPKULEENY 330
318 CRRNPQSESPWCCTTDPNIRVGYCSQIPNCDSH-----GODCYRGNG 360
331 CRRNPQSESPWCCTTTSSEVREWEYCT--IPSCSSPLSTERMDVVPPEQTPVPDQCHGNG 389
361 KNYMGLSQTGRSLGTCQWMDKNMEDJHRHFWME---PDASKULEENYCRNPDQAHQPMKY 417
390 QSYRGSSSTITIGRKQCSWSSMT--HRHKTPTENYPMAG--LTMNYCRNPDAD--KSPWCY 445
418 TGNPLIPMDYCPISRC 433
446 TTDPRVRYWEFCNLKCC 461

RESULT 14
PLHU

plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [misname]
N:Contents: angiotensin; microplasmin; plasminogen
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence, revision 02-Dec-1994 #text change 09-Jul-2004
C:Accession: A35229; 152242; A26646; 162738; 184609; S03735; A00929; A04627; A04625; A0
J. Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
R. Biol. Chem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fib
A:Reference number: A35229; MUID:90202879; PMID:2318848
A:Accession: A35229
A:Molecule type: DNA
A:Residues: 1-810 <PET>
A:Cross-references: UNIPROT:P00747; UNIPROT:Q9UBQ9; UNIPROT:Q9UMI2; GB:J05286; GB:M3427
A:Experimental source: leukocyte; lung fibroblast
R:Margaretti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; T
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A:Title: Definition of the transcription initiation site of human plasminogen gene in 1
A:Reference number: 152242; MUID:91097523; PMID:2268308
A:Accession: 152242
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <MAL>
A:Cross-references: GB:M62890; NID:9190092; PIDN:AAA36454.1; PID:g553613
R:Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human p
A:Reference number: A26646; MUID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471, 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:935530; PIDN:CAA28831.1; PID:g35531
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A:Reference number: 145961; MUID:85023311; PMID:6148961
A:Accession: 162738
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-471, 'D', 473-810 <MAL2>
A:Cross-references: GB:K02922; NID:9190112; PIDN:AAA60124.1; PID:g387031
A:Accession: 184609
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:K02921; NID:9190110; PIDN:AAA60123.1; PID:g190111
R:Brumsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1991
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03735
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOR>
R:Brumsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
submitted to the Atlas, July 1977
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOR>
R:Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; MUID:77225245; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <W1>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human pl;


```

Db      153 RNFNDNPQGWCTTTPERKYDYCDILECEE-ECMHCSGNNYDGKISKTMSSGLEQCAWDS 211
QY      202 QTPHRKFTLPERYPDGFDDNYCRNPDGQRPWCYTLDPHTTRREYCAIKTCADNTMNDUD 261
Db      212 QSPHAGYIPSKPKNLKKNYCRNPDRELPRWCFTTDPKRELCDIPRCT-----TP 265
QY      262 VPLE--TTECIQGGEGYRGVTIWNGLPCQWRDSQYPHEHDMTPENFKCKDLRENYCR 319
Db      266 PSSSGPTYQCLKGTGENYGNVAVTVSGHTCQWMSAQTPHTNRTPENFPCXKLDENYCR 325
QY      320 NPDGSESPWCFTTDPNIRGYCSQIPNCDSH-----GQDCYRGNKN 362
Db      326 NPDGKAPWCHTTNSQVRWEYC-KIPSCDSSPVSTEQLAPTAPPELTPVQDCYHGDQS 384
QY      363 YMGNLQTSRSLGTCSMWMDKNMEDLHRHIFWE---PDASKLENYCRNPDDDAHGFWCYTG 419
Db      385 YRGTSSTTTTGKKCCQSMSSWTP--HRHOKTPENYPNAG-LTMNYCRNPDAD-KGFWCFTT 440
QY      420 NPLIPWDYCPISRCEGDTTPTIV 442
Db      441 DSVRWREYCNLKKCSG-TEASV 462

```

RESULT 15

```

T18518
apolipoprotein(a) - western European hedgehog (fragment)
C:Species: Erinaceus europaeus (western European hedgehog)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18518
R:Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprot
A:Reference number: I46259; MID:96025778; PMID:7592597
A:Accession: T18518
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2869 <LAI>
A:Cross-references: UNIPROT:Q28398; EMBL:U31170; NID:g1046358; PID:g1046359; PIDN:AAC485
A:Experimental source: liver
C:Comment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, con
ent apolipoprotein(a).

```

Query Match 35.9%; Score 927; DB 2; Length 2869;
 Best Local Similarity 45.4%; Pred. No. 4.9e-58;
 Matches 172; Conservative 50; Mismatches 123; Indels 34; Gaps 11;

```

QY      89 ENKDYTRNCIIIGKRSYKGTVSTTKSGIKQWSSMI PHEHS-----YRGKDLQENYCRN 143
Db      2490 EPLDLVDQCLEGTGENYRGMAVTASGHTCQWRREQSPHSHSRTPENYPTKNLVGNVCRN 2549
QY      144 PRGEBGSPWCFTSNPEVRVEVCDIPQCS-----VECMTCNGESYRGLMDHTES 192
Db      2550 PGEV-APWCYTTNSAMRWEYCSIPACESPTPTTEHLVVEQCLENGENYQGNMAITVS 2608
QY      193 GKICQRMWDQTPHRHKLPERYPDKGFDDNYCRNPDGQRPWCYTLDPHTTRREYCAIKTC 252
Db      2609 GQPCQGMKQTPHRHKLPERYPDKGFDDNYCRNPDGQRPWCYTLDPHTTRREYCAIKTC 2668
QY      253 ADNTMNDTVPLETTTCIQGGEGYRGVTIWNGLPCQWRDSQYPHEHDMTPENFKCKD 312
Db      2669 -ESSSPTEPMIIPDQCLEGTGENYRGMAVTASGHTCQWRREQSPHSHSRTPENYPTKN 2727
QY      313 LRENYCRNPDGSESPWCFTTDPNIRGYCSQIPNCDSH-----GQDCYRGNKN 364
Db      2728 LRENYCRNPDGSESPWCFTTDPNIRGYCSQIPNCDSH-----GQDCYRGNKN 364
QY      365 GNLQTSRSLGTCSMWMDKNMEDLHRHIFWEBD--ASKLENYCRNPDDDAHGFWCYTGNP 421
Db      2787 GNLQTSRSLGTCSMWMDKNMEDLHRHIFWEBD--ASKLENYCRNPDDDAHGFWCYTGNP 421
QY      422 LIPWDYCPISRCEGDTTPT 440

```

Db 2843 AVRWEYCSIPTCESPTPT 2861
 Search completed: October 3, 2005, 05:59:59
 Job time : 39.775 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:29 ; Search time 64.6344 Seconds
(without alignments)
3501.834 Million cell updates/sec

Title: US-09-674-377b-2

Perfect score: 2579
Sequence: 1 ERKRNTHIEFKSAKATLTI.....IPDYCPISRCBGTPTIV 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2563.5	99.4	728	1 HGF_HUMAN	P14210 homo sapien
2	2424.5	94.0	730	2 Q867B7	Q867B7 canis fam1
3	2418.5	93.8	730	2 Q76BS1	Q76BS1 bos taurus
4	2414.5	93.6	728	2 Q9HH09	Q9HH09 felis silve
5	2371.5	92.0	728	2 Q8C9G5	Q8C9G5 mus musculu
6	2370.5	91.9	728	1 HGF_MOUSE	Q08048 mus musculu
7	2363.5	91.6	728	1 HGF_RAT	P17945 rattus norv
8	2014.5	78.1	726	2 Q90978	Q90978 gallus gall
9	1830	71.0	710	2 Q91402	Q91402 xenopus. he
10	1718	66.6	405	2 Q788Q2	Q788Q2 gallus gall
11	1204.5	46.7	449	2 Q6GP14	Q6GP14 xenopus lae
12	1198.5	45.9	716	2 Q91691	Q91691 xenopus lae
13	1184	45.4	704	2 Q90865	Q90865 gallus gall
14	1171	45.4	711	1 HGF_HUMAN	P26927 homo sapien
15	1163.5	45.1	717	2 P70006	P70006 xenopus lae
16	1156	44.8	709	2 Q7ZTN9	Q7ZTN9 xenopus lae
17	1152.5	44.7	709	2 Q90ZNE	Q90ZNE brachydantio
18	1151	44.6	567	2 Q13R08	Q13R08 homo sapien
19	1133.5	44.0	716	2 P70521	P70521 rattus norv
20	1114.5	43.2	716	1 HGF_MOUSE	P26928 mus musculu
21	1114.5	43.2	716	2 Q91XG8	Q91XG8 mus musculu
22	1073.5	41.6	818	2 Q6PBA6	Q6PBA6 brachydantio
23	1066	41.3	812	1 PLMN_RAT	Q01177 rattus norv
24	1056	40.9	806	1 PLMN_MACCU	Q18783 macropus eu
25	1054.5	40.9	790	1 PLMN_PIG	P06867 sus scrofa
26	1044.5	40.5	466	2 Q6TC10	Q6TC10 mus musculu
27	1044.5	40.3	812	1 PLMN_MOUSE	P20918 mus musculu
28	1036	40.2	215	2 O42341	O42341 gallus gall
29	1036	40.2	810	1 PLMN_ERIEU	Q29485 erinaceus e
30	1021	39.6	810	1 PLMN_MACMU	P12345 macaca mula
31	1019.5	39.5	812	1 PLMN_BOVIN	P06868 bos taurus

32	1018.5	39.5	810	1 PLMN_HUMAN	P00747 homo sapien
33	1013	39.3	759	2 Q7P84	Q7P84 rattus norv
34	971.5	37.7	359	2 Q8WNR1	Q8WNR1 canis fam1
35	958	37.1	648	2 Q9H1V4	Q9H1V4 homo sapien
36	927	35.9	2869	2 Q28398	Q28398 erinaceus e
37	864.5	33.5	4548	1 APOA_HUMAN	P08519 homo sapien
38	858.5	33.3	211	2 Q55027	Q55027 mus musculu
39	840	32.6	1420	1 APOA_MACMU	P14417 macaca mula
40	617	23.9	162	2 Q8C4E2	Q8C4E2 mus musculu
41	577.5	22.4	109	2 Q9N1B8	Q9N1B8 ovis aries
42	444.5	17.2	454	2 Q45506	Q45506 papio hamad
43	439	17.0	75	2 Q9BGN9	Q9BGN9 bos taurus
44	406	15.7	325	1 PLMN_PETMA	P33574 petromyzon
45	400	15.5	429	2 Q8AVB0	Q8AVB0 brachydantio

ALIGNMENTS

RESULT 1	HGF_HUMAN	STANDARD:	PRT: 728 AA.
AC	P14210; Q02935; Q13494; Q14519; Q8TCE2; Q9BYU9; Q9BYM0; Q9UDU6;		
DT	01-JAN-1990 (Rel. 13, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	25-OCT-2004 (Rel. 45, Last annotation update)		
DE	Hepatocyte growth factor precursor (Scatter factor) (SF)		
DE	(Hepatopoietin A).		
GN	Name:HGF; Synonyms:HPTA;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RX	MEDLINE=91340155; PubMed=1831432; DOI=10.1016/0378-1119(91)90080-U;		
RA	Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;		
RT	"Organization of the human hepatocyte growth factor-encoding gene.";		
RL	Gene 102:213-219 (1991).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	TISSUE=Placenta;		
RX	MEDLINE=89392017; PubMed=2528952;		
RA	Miyazawa K., Teubouchi H., Naka D., Takahashi K., Okigaki M.,		
RA	Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,		
RA	Gohda E., Daikuhara Y., Kitamura N.;		
RT	"Molecular cloning and sequence analysis of cDNA for human hepatocyte		
RT	growth factor.";		
RL	Biochem. Biophys. Res. Commun. 163:967-973 (1989).		
RN	[3]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	TISSUE=Leukocyte;		
RX	MEDLINE=91025062; PubMed=2145836;		
RA	Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T., Asami O.,		
RA	Hagiya M., Nakamura T., Shimizu S.;		
RT	"Isolation and expression of cDNA for different forms of hepatocyte		
RT	growth factor from human leukocyte.";		
RL	Biochem. Biophys. Res. Commun. 172:321-327 (1990).		
RN	[4]		
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 55-73 AND 495-520.		
RC	TISSUE=Liver;		
RX	MEDLINE=90066676; PubMed=2511289; DOI=10.1038/342440a0;		
RA	Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,		
RA	Sugimura A., Tashiro K., Shimizu S.;		
RT	"Molecular cloning and expression of human hepatocyte growth factor.";		
RL	Nature 342:440-443 (1989).		
RN	[5]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	TISSUE=Placenta;		
RX	MEDLINE=91200041; PubMed=1826653;		
RA	Miyazawa K., Kitamura A., Naka D., Kitamura N.;		
RT	"An alternatively processed mRNA generated from human hepatocyte		
RT	growth factor gene.";		

RL Eur. J. Biochem. 197;15-22(1991).
 RN [6]
 RN SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 583-592.
 RC TISSUE=lung fibroblast;
 RX PubMed=1824873;
 RA Rubin J.S., Chan A.M.-L., Bottaro D.P., Burgess W.H., Taylor W.G.,
 RA Cech A.C., Hirschfield D.W., Wong J., Miki T., Finch P.W.,
 RA Aaronson S.A.;
 RT "A broad-spectrum human lung fibroblast-derived mitogen is a variant
 RT of hepatocyte growth factor";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:415-419(1991).
 RN [7]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=embryonic fibroblast;
 RX MEDLINE=91334393; PubMed=1831266;
 RA Weidner K.M., Arakaki N., Hartmann G., Vandekeekhoeve J., Weingart S.,
 RA Riederer H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
 RA Birnmeier W.;
 RT "Evidence for the identity of human scatter factor and human
 RT hepatocyte growth factor";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
 RN [8]
 RN SEQUENCE FROM N.A. (ISOFORM 4).
 RX PubMed=1720571;
 RA Chan A.M.-L., Rubin J.S., Bottaro D.P., Hirschfield D.W., Chedid M.,
 RA Aaronson S.A.;
 RT "Identification of a competitive HGF antagonist encoded by an
 RT alternative transcript";
 RL Science 254:1382-1385(1991).
 RN [9]
 RN SEQUENCE FROM N.A. (ISOFORM 2), AND MUTAGENESIS OF ARG-494.
 RX MEDLINE=93087571; PubMed=1260830;
 RA Hartmann G., Naldini L., Weidner K.M., Sachs M., Vigna E.,
 RA Comoglio P.M., Birnmeier W.;
 RT "A functional domain in the heavy chain of scatter factor/hepatocyte
 RT growth factor binds the c-Met receptor and induces cell dissociation
 RT but not mitogenesis";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11574-11578(1992).
 RN [10]
 RN SEQUENCE FROM N.A. (ISOFORM 6).
 RX MEDLINE=96278713; PubMed=8662798; DOI=10.1074/jbc.271.22.13110;
 RA Clode V., Csaky K.G., Chan A.M.-L., Bottaro D.P., Taylor W.G.,
 RA Jansen R., Aaronson S.A., Rubin J.S.;
 RT "Hepatocyte growth factor (HGF)/NK1 is a naturally occurring
 RT HGF/scatter factor variant with partial agonist/antagonist activity";
 RL J. Biol. Chem. 271:13110-13115(1996).
 RN [11]
 RN SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS LYS-304 AND TYR-330.
 RA Riederer H., Livingston R.J., Daniels M.R., Chung M.-W.,
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
 RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
 RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department
 RT of Genome Sciences, Seattle, WA (URL: <http://egp.gs.washington.edu>).";
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBS databases.
 RN [12]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
 RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
 RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
 RA Wyllie G.A., Sekhon M., Becker M.C., O'Laughlin M.W., Schaller M.E.,
 RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Cordes M., Du H.,
 RA VanHunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
 RA Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
 RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
 RA Kozlowski-Feilly A., Leonard S., Kohling T., Rock S.M.,
 RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Stromatt C.,
 RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
 RA Wendt M.C., Yang S.-P., Schultz B.R., Wallis J.W., Splich J.,
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohlmann P.E., Cook L.L.,
 RA Hickenbotham M.T., Eldred J., Williams D., Bedell U.A., Mardis E.R.,
 RA Clifton S.W., Chisoe S.L., Marra M.A., Raymond C., Haugen E.,
 RA Gillett W., Zhou Y., James R., Phelps K., Iadonato S., Bubh K.,
 RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
 RA Baertsch R.A., Brent M.R., Keibler E., Flieck P., Bock P., Suyama M.,
 RA Bailey J.A., Portnoy M.R., Torrents D., Chinwalla A.T., Gish W.R.,
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
 RA Waterston R.H., Wilson R.K.;
 RT "The DNA sequence of human chromosome 7.";
 RL Nature 424:157-164(2003).
 RN [13]
 RN SEQUENCE FROM N.A. (ISOFORMS 5 AND 6).
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marutina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueidi T.B., Toshiyuki S., Cantini P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [14]
 RN SEQUENCE OF 1-208 AND 249-695 FROM N.A. (ISOFORM 1).
 RX MEDLINE=91369928; PubMed=1832556;
 RA Miyazawa K., Kitamura A., Kitamura N.;
 RT "Structural organization and the transcription initiation site of the
 RT human hepatocyte growth factor gene";
 RL Biochemistry 30:9170-9176(1991).
 RN [15]
 RN SIGNAL SEQUENCE CLEAVAGE SITE.
 RX MEDLINE=91207365; PubMed=1826837;
 RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirose S., Kondo J.,
 RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
 RA Hishida T., Daikuhara Y.;
 RT "Identification of the N-terminal residue of the heavy chain of both
 RT native and recombinant human hepatocyte growth factor";
 RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
 RN [16]
 RN CARBOHYDRATE-LINKAGE SITE. THR-476.
 RX MEDLINE=93129192; PubMed=1482348;
 RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
 RA Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 RT on the alpha chain";
 RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 RN [17]
 RN MUTAGENESIS.
 RX MEDLINE=92331602; PubMed=1321034;
 RA Loker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 RA Baker J.B., Godowski P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 RT identification of variants that lack mitogenic activity yet retain
 RT high affinity receptor binding";
 RL EMBO J. 11:2503-2510(1992).
 RN [18]
 RN STRUCTURE BY NMR OF 31-127.
 RX MEDLINE=98154523; PubMed=9493272; DOI=10.1016/S0969-2126(98)00012-4;
 RA Zhou H., Mazulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
 RA Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 RT factor reveals a potential heparin-binding site";
 RL Structure 6:109-116(1998).
 RN [19]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.

Query Match	99.4%	Score 2563.5	DB 1	Length 728
Best Local Similarity	98.7%	Pred. No. 2e-185		
Matches 441	Conservative 1	Mismatches 0	Indels 5	Gaps 1
QY	1	ERKRRTTHIEFKKSAKTLIKIDPALIKITKKVNTADQCANRCTRNKGLPFTCAFAVBD	60	
Db	32	QKRKRRTTHIEFKKSAKTLIKIDPALIKITKKVNTADQCANRCTRNKGLPFTCAFAVBD	91	
QY	61	ARKQCLMFPNMSMGVKKERFGEHEDLYENKDYIRNCLIGGRSYKGVSTITKGIKQP	120	
Db	92	ARKQCLMFPNMSMGVKKERFGEHEDLYENKDYIRNCLIGGRSYKGVSTITKGIKQP	151	
QY	121	WSSMIPHEH-----SYRKDIQENYCRNPRGEGSPMCFSTNSPERYEVCQIPQCEVEYC	175	
Db	152	WSSMIPHEHSLFPLSSRYGRKDIQENYCRNPRGEGSPMCFSTNSPERYEVCQIPQCEVEYC	211	
QY	176	MTCNGESYRGLMDHTSGSKICQRMWDHQTPHRHKFLPERYPDKGPDNDYCRNPDQPRWC	235	
Db	212	MTCNGESYRGLMDHTSGSKICQRMWDHQTPHRHKFLPERYPDKGPDNDYCRNPDQPRWC	271	
QY	236	YTLDPHTFWEYCAITTCADNTMNDTDLVLETTETCIQGGEGYRGTVNTIANGIPQQRMS	295	
Db	272	YTLDPHTFWEYCAITTCADNTMNDTDLVLETTETCIQGGEGYRGTVNTIANGIPQQRMS	331	
QY	296	QYPHEHDMTPPNPFCKDKLRENYCRNPDGSESPMCFTTDPNIRVGCQOIPNCDMSHGDC	355	
Db	332	QYPHEHDMTPPNPFCKDKLRENYCRNPDGSESPMCFTTDPNIRVGCQOIPNCDMSHGDC	391	
QY	356	YRGNKXNMGSLQSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDADHGPW	415	
Db	392	YRGNKXNMGSLQSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDADHGPW	451	
QY	416	CYTGNPLIPWDYCPISRCGDTTPTIV	442	
Db	452	CYTGNPLIPWDYCPISRCGDTTPTIV	478	

RESULT 2

Q867B7 PRELIMINARY; PRT; 730 AA.

AC Q867B7;

DT 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE Hepatocyte growth factor.

GN Name=HGF;

OS Carls familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

 (1)

RN SEQUENCE FROM N.A.

RP STRAIN=Beagle;

RC MEDLINE=22844761; PubMed=12963274; DOI=10.1016/S0165-2427(03)00118-1;

RX Miyake M., Saze K., Yaguchi T., Wang J., Suzuta Y., Haga Y.,

RA Takahashi S.Y., Yamamoto Y., Iwabuchi S.;

RT "Canine hepatocyte growth factor: molecular cloning and

RT characterization of the recombinant protein.";

RL Vet. Immunol. Immunopathol. 95:135-143(2003).

 (2)

RP SEQUENCE FROM N.A.

RL Liao A.T., Chien M.B., London C.A.;

RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: Belongs to peptidase family 11.

DR EMBL: AB090353; BACS7560.1; -

DR EMBL: AY543632; AAS48570.1; -

DR HSSP: P14210; IGMN.

DR MEROPS: S01.982; -

DR GO: GO:0004263; F:chymotrypsin activity; IEA.

DR GO: GO:0008233; F:peptidase activity; IEA.

[illegible]

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamaji D., Kimura K., Matanabe A., Makondo K., Saito M.,
 RL Submitted (May-2003) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 DR EMBL: AB110822; BAD02475.1; -.
 DR HSSP: P08519; IKTV.
 DR GO: GO:0004263; F:chymotrypsin activity; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR011359; HGF_MSTL.
 DR InterPro: IPR000001; Kirtingle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR Pfam: PF00051; Kirtingle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; Trypsin; 1.
 DR PIRSF: PIRSF001152; HGF_MSTL; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRODOM: PD000395; Kirtingle; 4.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00021; KRINGLE_1; 4.
 DR PROSITE: PS50070; KRINGLE_2; 4.
 DR PROSITE: PS50948; PAN; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR Hydrolase: Kirtingle; Protease; Serine protease.
 KW Hydrolyse, Kirtingle, protease, Serine protease.
 SQ SEQUENCE 730 AA; 83357 MW; E39D653B0A85F49B CRC64;

Query Match 93.8%; Score 2418.5; DB 2; Length 730;
 Best Local Similarity 91.5%; Pred. No. 1.8e-174;
 Matches 409; Conservative 22; Mismatches 11; Indels 5; Gaps 1;

QY 1 ERKRRNTLHEPKSAKTLIKIDPLKIKTKKVTADQCANRCTRNKGLPTCKAFVDPK 60
 DB 32 QKKRRNTLHEPKSAKTLIKIDPLKIKTKKVTADQCANCTRNKGLPTCKAFVDPK 91

QY 61 ARKQCLWFPFNSMSGVKKFGEHFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCP 120
 DB 92 ARKRCILWFPFNSMSGVKKFGEHFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCP 151

QY 121 WSSMTLPHH-----SYRGKDLQENYCRNRGBEGGPMCTSNPEVRYEVCIDIPQSEVVC 175
 DB 152 WSSMTLPHHESFLPSSYRGKDLQENYCRNRGBEGGPMCTSNPEVRYEVCIDIPQSEVVC 211

QY 176 MTCNGESYRGKDLQENYCRNRGBEGGPMCTSNPEVRYEVCIDIPQSEVVC 235
 DB 212 MTCNGESYRGKDLQENYCRNRGBEGGPMCTSNPEVRYEVCIDIPQSEVVC 271

QY 236 YLLDHTHTREYCAIKTCANTMADVDVLETTETCLOGEGSIRGYNTIWNIGIPQQRWMS 295
 DB 272 YLLDHTHTREYCAIKTCANTMADVDVLETTETCLOGEGSIRGYNTIWNIGIPQQRWMS 331

QY 332 QIPPHQDITPENFKCDLENYCRNPDGSESPMCTTDNINIRGYCSQIPNCDMSHGDC 355
 DB 332 QIPPHQDITPENFKCDLENYCRNPDGSESPMCTTDNINIRGYCSQIPNCDVSSGDC 391

QY 392 YRNGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASKLXENYCRNPDGDAHGPW 415
 DB 416 YRNGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASKLXENYCRNPDGDAHGPW 451

QY 416 YRNGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASKLXENYCRNPDGDAHGPW 451

DB 452 CYTGNPLIPWDYCPISRCGSDTPTIV 478

RESULT 4
 ID Q9BH09 PRELIMINARY; PRT; 728 AA.
 AC Q9BH09;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Hepatocyte growth factor HGF.
 GN Name:HGF;
 OS Fells silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felle.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Kobayashi Y., Nakamura N., Ishizaka T., Maeda K., Ohno K.,
 RA Tsujimoto H.;
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Miyake M., Yaguchi T., Saze K., Suzuta Y., Wang J., Okazaki M.,
 RA Haga Y., Yamamoto Y., Takahashi S., Iwabuchi S.;
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 DR EMBL: AB046610; BAB21499.1; -.
 DR EMBL: AB080187; BAC10545.1; -.
 DR HSSP: P14210; IGMN.
 DR MEROPS: S01.982; -.
 DR GO: GO:0004263; F:chymotrypsin activity; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR011359; HGF_MSTL.
 DR InterPro: IPR000001; Kirtingle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR Pfam: PF00051; Kirtingle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR PIRSF: PIRSF001152; HGF_MSTL; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRODOM: PD000395; Kirtingle; 4.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00021; KRINGLE_1; 4.
 DR PROSITE: PS50070; KRINGLE_2; 4.
 DR PROSITE: PS50948; PAN; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR Hydrolase: Kirtingle; Protease; Serine protease.
 KW Hydrolyse, Kirtingle, protease, Serine protease.
 SQ SEQUENCE 728 AA; 807F4A33301B190A CRC64;

Query Match 93.6%; Score 2414.5; DB 2; Length 728;
 Best Local Similarity 91.5%; Pred. No. 3.7e-174;
 Matches 409; Conservative 21; Mismatches 12; Indels 5; Gaps 1;

QY 1 ERKRRNTLHEPKSAKTLIKIDPLKIKTKKVTADQCANRCTRNKGLPTCKAFVDPK 60
 DB 30 QKKRRNTLHEPKSAKTLIKIDPLKIKTKKVTADQCANCTRNKGLPTCKAFVDPK 89

QY 61 ARKQCLWFPFNSMSGVKKFGEHFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCP 120
 DB 90 ARKRCILWFPFNSMSGVKKFGEHFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCP 149

QY 121 WSSMTLPHH-----SYRGKDLQENYCRNRGBEGGPMCTSNPEVRYEVCIDIPQSEVVC 175


```
Db 150 WNSMIPHEHSFLPSSIRGKLOEYCNPRGEGGPMCFPSNBEVREVCDDIQCESEVC
Qy 176 MTCNGESYRGIMDHTESSGKICORWDHQPPIRHHKFLPERYPDKGFDDNYCNPDPGPPWC
Db 210 MTCNGESYRPMHHTSSGKICORWDHQPPIRHHKFLPERYPDKGFDDNYCNPDPGPPWC
Qy 236 YTLDPHTRMWYCAIKTCADNTAMDTPVLETTETCICQGGEGYRGVTNTWNGIPCORWDS
Db 270 YTLDPHTRMWYCAIKTCADNTAMDTPVLETTETCICQGGEGYRGVTNTWNGIPCORWDS
Qy 296 QYEHEDMPENKRCOLRENYCNRNPDGSEPMCFPTDPIRIGYCSQIPNCMSHGDC 355
Db 330 QYEHEDMPENKRCOLRENYCNRNPDGSEPMCFPTDPIRIGYCSQIPNCMSHGDC 389
Qy 356 YRGNGKNYKGNLTSQTSGLTCSMMWDMEDLHRIHIEWEPASKLNEYCNPDDAHGPM 415
Db 390 YRGNGKNYKGNLTSQTSGLTCSMMWDMEDLHRIHIEWEPASKLNEYCNPDDAHGPM 449
Qy 416 CYTGNPLIPMDYCPISRCGEGDTPTIV 442
Db 450 CYTGNPLIPMDYCPISRCGEGDTPTIV 476

RESULT 5
ID 08C9G5 PRELIMINARY; PRT; 728 AA.
AC 08C9G5;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
DE library, clone:A630060C16 product:hepatocyte growth factor, full
DE insert sequence.
GN Name:Hgf;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
```

```
RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Katsunai T., Taahito H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujitake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai U., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK042121; BAC31175.1; -.
DR HSSP: P14210; 1GMN.
DR MCD; MG1:96079; Hgf.
DR GO; GO:0008283; P:cell proliferation; IDA.
DR GO; GO:0009092; P:cellular morphogenesis; IDA.
DR InterPro; IPR011359; HGF_MST1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PIRSF001152; HGF_MST1; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00948; PAN; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
KW Hydroxylase, Kringle, Protease, Serine protease.
SQ SEQUENCE 728 AA; 82990 MW; 8D5258DF3BCF3545 CRC64;

Query Match 92.0%; Score 2371.5; DB 2; Length 728;
Best local Similarity 89.9%; Pred No. 6 6e-17;
Matches 402; Conservative 21; Mismatches 19; Indels 5; Gaps 1;
```

```
Qy 1 ERKRNTLHEFKSAKTLTIKIDPALKI KTKKVTNADQCANCTRNGKLPFTCKAFVDPK 60
Db 33 OKRRNTLHEFKSAKTLTIKEDPLKI KTKKVNAGADECANCRINRGTFTCKAFVDPK 92
Qy 61 AKQCIMPFPNMSGYKKEFGHEFDLYENKQYINRCIIIGKRSTKGYSTIRKSGIKCOP 120
Db 93 SKRCYWPFPNMSGYKKEFGHEFDLYENKQYINRCIIIGKSGYKGYSTIRKSGIKCOP 152
Qy 121 WNSMIPHEH-----SYRGKLOEYCNPRGEGGPMCFPSNBEVREVCDDIQCESEVC 175
Db 153 WNSMIPHEHSFLPSSIRGKLOEYCNPRGEGGPMCFPSNBEVREVCDDIQCESEVC 212
Qy 176 MTCNGESYRGIMDHTESSGKICORWDHQPPIRHHKFLPERYPDKGFDDNYCNPDPGPPWC 235
```

Db 213 MTCNSESRYGPMWHTSGKTCQRMDOQTPEHRKFLPERYPDXGFPDNYCRNPDGKPRPWC 272
Qy 226 YTLDDHTWMEVCAITTCADNTMNDVDVPLETTECIQOGGEGRGVNTMNIICORWDS 295
Db 273 YTLDDHTWMEVCAITTCADNTMNDVDVPLETTECIQOGGEGRGVNTMNIICORWDS 332
Qy 296 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTTPNIRVGYCSQIPIPCDMSHGDC 355
Db 333 QYPHHDITPENFKCKDLRENYCRNPDGSESPWCTTTPNIRVGYCSQIPIPCDMSHGDC 392
Qy 356 YRGNKNTYMGUSQTRSGITCSMDNKNMEDLHRHIFWEPDASKLKNYCRNPDGDAHGPW 415
Db 393 YRGNKNTYMGUSQTRSGITCSMDNKNMEDLHRHIFWEPDASKLKNYCRNPDGDAHGPW 452
Qy 416 CYTGNPLIPWDYCPISRCGSDTPTTV 442
Db 453 CYTGNPLIPWDYCPISRCGSDTPTTV 479

RESULT 6

HGF_MOUSE
ID_HGF_MOUSE STANDARD; PRT; 728 AA.
AC Q08048; Q61662; Q64007;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
GN (Hepatopoietin A).
OS Name=Hgf;
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), AND SEQUENCE OF 496-504.
RX TISSUE=Mammary fibroblast;
RA MEDLINE=9418357; PubMed=815822;
RA Sasaki M., Nishio M., Sasaki T., Enami J.;
RT "Identification of mouse mammary fibroblast-derived mammary growth
factor as hepatocyte growth factor.";
RL Biochem. Biophys. Res. Commun. 199;779-779(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA MEDLINE=94363381; PubMed=8081873;
RA Lee C.C., Kozak C.A., Yamada K.M.;
RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter
factor gene.";
RL Cell Adhes. Commun. 1;101-111(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA MEDLINE=94060105; PubMed=8241272; DOI=10.1016/0167-4781(93)90159-B;
RA Liu Y., Michalopoulos G.K., Zarnegar R.;
RT "Molecular cloning and characterization of cDNA encoding mouse
hepatocyte growth factor.";
RL Biochim. Biophys. Acta 1216;299-303(1993).
CC -|- FUNCTION: HGF is a potent mitogen for mature parenchymal
hepatocyte cells, seems to be an hepatotrophic factor, and acts as
growth factor for a broad spectrum of tissues and cell types. It
has no detectable protease activity.
CC -|- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
disulfide bond.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q08048-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q08048-2; Sequence=VSP_005408;
CC -|- SIMILARITY: Belongs to the peptidase S1 family. Plasminogen
subfamily.
CC -|- SIMILARITY: Contains 4 kringle domains.

CC -|- SIMILARITY: Contains 1 PAN domain.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

CC -----
CC EMBL: D10212; BAA01064.1; -;
CC EMBL: D10213; BAA01065.1; -;
CC EMBL: S71816; BAB31855.1; -;
CC EMBL: X72307; CAA51054.1; ALT_INIT.
CC PIR: JC2117; A60185.
CC HSP: P14210; IBBT.
CC MEROPS: S01.982; -;
CC MGD; MGI:96079; HGF.
CC GO; GO:0008283; P:cell proliferation; IDA.
CC GO; GO:000902; P:cellular morphogenesis; IDA.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR003014; PAN.
CC InterPro: IPR003609; Pan_app.
CC InterPro: IPR009003; pep_Ser_Cys.
CC InterPro: IPR001254; peptidase_S1.
CC Pfam: PF00051; Kringle; 4.
CC Pfam: PF00024; PAN; 1.
CC Pfam: PF00089; Trypsin; 1.
CC PIRSF: PIRSF001152; HGF_MST1; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRODOM: PD000395; Kringle; 4.
CC SMART: SM00130; KR; 4.
CC SMART: SM00473; PAN_AP; 1.
CC SMART: SM00020; TRYP_Ser; 1.
CC PROSITE: PS00021; KRINGLE_1; 4.
CC PROSITE: PS50070; KRINGLE_2; 4.
CC PROSITE: PS50240; TRYPSIN_DOM; 1.
CC Alternative splicing: Direct protein sequencing; Glycoprotein;
CC Growth factor; Kringle; Pyroglutamate carboxylic acid; Repeat;
CC Serine protease homology; Signal.
CC SIGNAL 1 32
CC CHAIN 33 495
CC CHAIN 496 728
CC MOD_RES 33 33
CC DOMAIN 38 124
CC DOMAIN 129 207
CC DOMAIN 212 289
CC DOMAIN 306 384
CC DOMAIN 392 470
CC DOMAIN 496 728
CC DISULFID 71 97
CC DISULFID 75 85
CC DISULFID 488 607
CC CARBOHYD 295 295
CC CARBOHYD 403 403
CC CARBOHYD 569 569
CC CARBOHYD 656 656
CC VARSPPLIC 163 167
CC CONFLICT 344 344
CC CONFLICT 479 479
CC CONFLICT 564 564
CC SEQUENCE 728 AA; 82944 MW; A0381FC497534328 CRC64;

Query Match 91.9%; Score 2370.5; DB 1; Length 728;
Best Local Similarity 89.9%; Pred. No. 7.8e-171;
Matches 402; Conservative 21; Mismatches 19; Indels 5; Gaps 1;
Qy 1 ERKRNTIHEPKSAKTLIKIDPALKIKTKKVTADQACNCRTRNKGLPFTCKAFVFDK 60

```
Db 33 QKRRNTLHEFKKSAATTLTKEDPLKIKTKKYNASDECANRRIIRNGPFTCKAFVFDK 92
Qy 61 ARKQCLMPFPNSNSGVKKFEGHEFDLYENKDYIRNCIIIGKSGYGTVAITSIGIKCOP 120
Db 93 SRKRCYWPFPNSNSGVKKFGHEFDLYENKDYIRNCIIIGKSGYGTVAITSIGIKCOP 152
Qy 121 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIQCSVEVC 175
Db 153 WNSMIDHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIQCSVEVC 212
Qy 176 MTCNGSSYRGKLMHTSSGKICQWHDQTPRHAKFLPERYDKGPDNYCNPFGCQRPWC 235
Db 213 MTCNGSSYRGKPMHTSGKICQWHDQTPRHAKFLPERYDKGPDNYCNPFGCQRPWC 272
Qy 236 YTLDPHTREYCAIKTCADNTAMDTPVPLETTECICQCGEGYGVNTIINGI.PCORWDS 295
Db 273 YTLDPHTREYCAIKTCADNTAMDTPVPLETTECICQCGEGYGVNTIINGI.PCORWDS 332
Qy 296 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDC 355
Db 333 QYPHKADITPENFKCDLRENYCRNPDGASPMCFSTTDPNIRVGYCSQIPKCDVSSGQDC 392
Qy 356 YRGNGKYNMGNTLSQTSGLTCSWMDQXMDLHRIHFEPPASTKLNEYCNPDDDAHGPW 415
Db 393 YRGNGKYNMGNTLSKTSGLTCSWMDQXMDLHRIHFEPPASTKLNEYCNPDDDAHGPW 452
Qy 416 CYTGNPLIPMDYCPISRCGDTPTTY 442
Db 453 CYTGNPLIPMDYCPISRCGDTPTTY 479
```

```
RESULT 7
HGF_RAT ID HGF_RAT STANDARD; PRT; 728 AA.
AC P17545;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN Name=Hgf;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Miscar; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshiro K., Hagiya M., Nishizawa T., Seki T., Shimomishi M.,
RA Shimizu S., Nakamura T.;
RT "Deduced primary structure of rat hepatocyte growth factor and
RT expression of the mRNA in rat tissues.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Miscar; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "Primary structure of rat hepatocyte growth factor and induction of
RT its mRNA during liver regeneration following hepatic injury.";
RL Eur. J. Biochem. 193:375-381(1990).
CC -|- FUNCTION: HGF is a potent mitogen for mature parenchymal
CC hepatocyte cells, seems to be an hepatotrophic factor, and acts as
CC growth factor for a broad spectrum of tissues and cell types. It
CC has no detectable protease activity.
CC -|- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
CC disulfide bond.
CC -|- SIMILARITY: Belongs to the peptidase S1 family. Plasminogen
CC subfamily.
CC -|- SIMILARITY: Contains 4 kringle domains.
CC -|- SIMILARITY: Contains 1 PAN domain.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
Db EMBL: D90102; BAA14133.1; -
Db EMBL: X54400; CA38266.1; -
Db PIR: A35644; A35644.
Db HSSP: P14210; 1BHT.
Db MEROPS: S01.978; -.
Db RGD: 2794; Hgf.
Dr InterPro: IPR000001; Kringle.
Dr InterPro: IPR003014; PAN.
Dr InterPro: IPR003609; Pan_app.
Dr InterPro: IPR009003; Pept_Ser_Cys.
Dr InterPro: IPR001254; Peptidase_S1.
Dr InterPro: IPR001314; Peptidase_S1A.
Dr Pfam: PF000051; Kringle_4.
Dr Pfam: PF00024; PAN_1.
Dr Pfam: PF00089; Trypsin_1.
Dr PIRSF: PIRSF001152; HGF_MST1_1.
Dr PRINTS: PR00722; CHYMOTRYPSIN.
Dr PRINTS: PR00018; KRINGLE.
Dr ProDom: PD000395; Kringle_4.
Dr SMART: SM00130; KR_4.
Dr SMART: SM00473; PAN_Ap_1.
Dr SMART: SM00020; Tryp_Spc_1.
Dr PROSITE: PS00021; KRINGLE_1; 4.
Dr PROSITE: PS50070; KRINGLE_2; 4.
Dr PROSITE: PS50948; PAN_1.
Dr PROSITE: PS50240; TRYPsin_DOM_1.
KW Direct protein sequencing; Glycoprotein; Growth factor; Kringle;
KW Pyroliidone carboxylic acid; Repeat; Serine protease homolog; Signal.
FT SIGNAL 1 32
FT CHAIN 33 495
FT CHAIN 496 728
FT MOD_RES 33 33
FT DOMAIN 38 124
FT DOMAIN 129 207
FT DOMAIN 212 289
FT DOMAIN 305 384
FT DOMAIN 392 470
FT DOMAIN 496 728
FT DISULFID 71 97
FT DISULFID 75 85
FT DISULFID 488 607
FT CARBOHYD 295 295
FT CARBOHYD 403 403
FT CARBOHYD 569 569
FT CARBOHYD 656 656
SQ SEQUENCE 728 AA; 82905 MW; 3E0BF1F96ADCEDFF CRC64;

Query Match 91.6%; Score 2363.5; DB 1; Length 728;
Best Local Similarity 89.3%; Pred. No. 2,6e-170;
Matches 399; Conservative 23; Mismatches 20; Indels 5; Gaps 1;

Qy 1 EKKRRNTLHEFKKSAATTLTKEDPLKIKTKKYNASDECANRRIIRNGPFTCKAFVFDK 60
Db 33 QKRRNTLHEFKKSAATTLTKEDPLKIKTKKYNASDECANRRIIRNGPFTCKAFVFDK 92
Qy 61 ARKQCLMPFPNSNSGVKKFEGHEFDLYENKDYIRNCIIIGKSGYGTVAITSIGIKCOP 120
Db 93 SRKRCYWPFPNSNSGVKKFGHEFDLYENKDYIRNCIIIGKSGYGTVAITSIGIKCOP 152
Qy 121 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIQCSVEVC 175
Db 153 WNSMIDHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIQCSVEVC 212
```

Qy	176	MTCNESYVGLMDHTHESGTCQRMWQHOTHRHFLPERYDDKGFDDNNCGNPDGQGRPMC	235
Db	213	MTCNESYVGLMDHTHESGTCQRMWQOTPHRHFLEPERYDDKGFDDNNCGNPDGKRPBWC	272
Qy	236	YTLDPHTRWVCAIKTCADNTMNDITDVPLETTECIGQGGEGVGRVNTIWNIGIPCORWDS	295
Db	273	YTLDPDTPWVCAIKKCAHSAVNETDVPMETTECIGKQGGVGRVNTIWNIGIPCORWDS	332
Qy	256	QYPHEDHMPPEMPKCDLSENYCRNPDGSESFPCFTTDPNIRVGYCSQIPNCMSHGQDC	355
Db	333	QYPHKHDTIPENPKCDLSENYCRNPDGASFPCTTDPNIRVGYCSQIPKCVVSSGQDC	392
Qy	356	YRGNGKNVYGNLSQTRSGLTCSMWDMKMDLHRHIFWEPDASKLNNENYCRNPDDAHGPW	415
Db	393	YRGNGKNVYGNLSKTRSGLTCSMWDMKMDLHRHIFWEPDASKLTNKNYCNRPDDDAHGPW	452
Qy	416	CYTGNDPLMPDYCPISRCEGDTPTTV	442
Db	453	CYTGNDPLVWDYCPISRCGGDTPTTV	479

RESULT 8

ID	Q90978.	PRELIMINARY;	PRT;	726 AA.
AC	Q90978.	Q90866;		
DT	01-JAN-1998	(TREMBLrel. 05. Created)		
DT	01-JAN-1998	(TREMBLrel. 05; Last sequence update)		
DT	01-MAR-2004	(TREMBLrel. 26; Last annotation update)		
DE	Hepatocyte growth factor /scatter factor.			
GN	Name=HGF/SF;			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
OC	Archosauirata; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
CC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96029010.; PubMed=7554499;			
RA	They C., Shape M.J., Batley S.J., Stern C.D., Cherard J.E. ;			
RT	"Expression of HGF/SF, HGFI/MSP and c-met suggests new functions			
RT	during early chick development.";			
RL	Dev. Genet. 17:90-101(1995).			
CC	-I- SIMILARITY: Belongs to peptidase family SI.			
DR	EMBL; X84045; CA858864.1; --.			
DR	PIR; I51285; I51285.			
DR	HSSP; P14210; IGMP.			
DR	GO; GO:0004263; F:chymotrypsin activity; IEA.			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	GO; GO:0004295; F:trypsin activity; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	Pfam; PF00051; Kingle; 4.			
DR	Pfam; PF00024; PAN; 1.			
DR	Pfam; PF00089; Trypsin; 1.			
DR	PIRSF; PIRSF001152; HGF_MST1; 1.			
DR	PRINTS; PRO0722; CHYMOTRYPSIN.			
DR	PRINTS; PRO0018; KINGLE.			
DR	ProDom; PD000395; Kingle; 4.			
DR	SMART; SM00130; KR; 4.			
DR	SMART; SM00473; PAN_AP; 1.			
DR	SMART; SM00020; TRYP_SPE; 1.			
DR	PROSITE; PS00021; KINGLE_1; 4.			
DR	PROSITE; PS50070; KINGLE_2; 4.			
DR	PROSITE; PS50948; PAN; 1.			
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.			
KW	Hydrolase; kingle; protease; Serine protease.			
SQ	SEQUENCE 726 AA; 82865 MW; 5BD06CFB5C40B003 CRC64;			

Query Match	78.1%;	Score 2014.5;	DB 2;	Length 726;
Best Local Similarity	75.5%;	Pred. No. 6,5e-144;		
Matches 336;	Conservative 49;	Mismatches 53;	Indels 7;	Gaps 2;
QY	3	KRRNTIHEKKSAKTTLLIKIDPALKTKTKVNTADOCANRCRNRKGLPTCAAFVDDKAR	62	

```

Db      30  XRRNLDHYKKTGELMLIKVNTLEVKTKLNTTSGCAKRGSRNGLSPTCAFAVDRVT  89
Qy      63  KOCIMFPNPNSSGVYKKEBGEFDLYENKDYRNCITGKGRSYKGTVITTKSGIKCQPM  122
Db      90  KRCHMLSNLSLNTNGRKQKODHAFDFLEKKDYRNCITGGAAYKGTISITKSGIOQAWN  149
Qy     123  SMIPHEH-----SYRGKDLQENYCNPNRGEEGPMCFTSNPEYREYVCIIPOCEVECM  177
Db     150  SMIPHEHGFPLPSYSGKDLRENYCNRNPRGEBGPMCFTTSPQMRHEVCIIPLCEVECM  209
Qy     178  CNGESYRGMLDHTESGKICQRMWHOTPHRKELPERYPDKGEFDNYCRNPDGQPRWCYT  237
Db     210  CNGESYRGMDHTESGKEQRMWDLQRPHKHKRPREYPRDKGFDNYCRNPDGKILPMCYT  269
Qy     238  LDPHTRWEYCAIKTADNTMNTDVPLETTETCIIQOGEGYRGTYNTIMGICPCQRMSOY  297
Db     270  LDPNTPWECAIKTQDVGILNSTEVAETTTICQOGEGYRGTYNTIMSGICQQRMSOF  329
Qy     298  PHEHMTPENFCKDLRENYCNRNPGSSEPMCFITDPNIRVGYCSOIPNCDSMHSQDCYR  357
Db     330  PHQHNTPENFCKDLRENYCNRNPGSSEPMCFITDPNIRIGYCSOIPKCDVSNEDCYR  389
Qy     358  GNGKNYMGULSOTRSGLTCSMWDKNMEDLHRH--IFWEBDASKLENYCRNPDDBAHGPW  415
Db     390  GNGKSYMGULSNTRIGLTCSYWDKNIEDLRRIQIFREBDVSKLKNKYCRNPDDBFHGPW  449
Qy     416  CYTGNPLIPWDYCPISRCEGDTTPT  440
Db     450  CYTDPLIPWDYCPISRCTGDTTPT  474

```

RESULT 9

ID	PRELIMINARY:	PRT:	710 AA.
AC	091402;		
DT	01-NOV-1996 (TREMBLrel. 01.	Created)	
DT	01-NOV-1996 (TREMBLrel. 01.	Last sequence update)	
DT	01-MAR-2004 (TREMBLrel. 26,	Last annotation update)	
DE	Hepatocyte growth factor.		
GN	Name=HGF;		
OS	Xenopus.		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;		
OC	Xenopodine.		
OX	NCBI_Taxid=8353;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Tailbud;		
RX	MEDLINE:95567990; PubMed=7748783; DOI=10.1016/0922-4773(94)00309-B;		
RA	Nakamura H., Tashiro K., Nakamura T., Shiohara K.;		
RT	Molecular cloning of Xenopus HGF cDNA and its expression studies in		
RL	Xenopus early embryogenesis.";		
CC	Mech. Dev. 49:123-131 (1995).		
CC	-I- SIMILARITY: Belongs to: peptidase family S1.		
DR	EMBL: S77422; AAB34354.2; -.		
DR	HSSP: P14210; IGMN.		
DR	MEROPS: S01.976; -.		
DR	GO: GO:0004263; F:chymotrypsin activity; IEA.		
DR	GO: GO:0008233; F:peptidase activity; IEA.		
DR	GO: GO:0004295; F:trypsin activity; IEA.		
DR	GO: GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	Pfam: PF00051; Kringle; 4.		
DR	Pfam: PF00024; PAN; 1.		
DR	Pfam: PF00089; Trypsin; 1.		
DR	PIRSF: PIRSF001152; HGF MS1; 1.		
DR	PRINTS: PR00722; CHYMOTRYPSIN.		
DR	PRINTS: PR00018; KRINGLE.		
DR	ProDom: PD000395; Kringle; 4.		
DR	SMART: SM00130; KR; 4.		
DR	SMART: SM00473; PAN_AP; 1.		
DR	SMART: SM00020; Tryp_Spc; 1.		
DR	PROSITE: PS00021; KRINGLE_1; 4.		
DR	PROSITE: PS00070; KRINGLE_2; 4.		

RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchan J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Klein S., Strausberg R.,
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC073334; AAH73334.1; -
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0003809; F:thrombin activity; IEA.
 DR GO; GO:0007596; P:blood coagulation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000001; Kringleg.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR003966; Peptidase_S1A_pr.
 DR Pfam; PF00051; Kringleg; 4.
 DR Pfam; PF00024; PAN; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR Prodom; PD000395; Kringleg; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN AP; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS0070; KRINGLE_2; 4.
 DR PROSITE; PS00948; PAN; 1.
 KW Kringleg.
 SQ SEQUENCE 449 AA; 52467 MW; 5C2CE4B708E13D0B CRC64;
 Query Match 46.7%; Score 1204.5; DB 2; Length 449;
 Best Local Similarity 47.0%; Pred. No. 6.7e-83;
 Matches 205; Conservative 74; Mismatches 144; Indels 13; Gaps 6;
 Oy 5 RNTIEFFKSAKTTLIKIDPALKIKTKKYNVTADOCANRCTRNKGLPFTCKAFVFDKARQ 64
 Db 21 RALNDYQSKGLIELVHMNG-GVKOEIOSEIOVCAKQSD----LLDCRSFYVMKSGS 75
 Oy 65 CLMFPNMSGGVKKEFGHEFDLYENKDYIRNCITIGKRSYKTVSYTSGIKQCPWSSM 124
 Db 76 CHLLPMTQNSANVLLQRNVQYDLYQKDYIRDCVANGNTYRGVSKTSKGRCCWRMLK 135
 Oy 125 IPHEHY--RGKDLQENYCRNPRGEGGPGWCTSNPEVRYEVCDIPOCSEVECMTCNGE 181
 Db 136 FPHDHKFSPIHWPDLLENYCRNPDSPGPGWCTTIDKNIHQYCGIKKCEDAVCLTCNGE 195
 Oy 182 SYRGLMDHTESGKICQRMWDQTPHRRKFLPERYPDKGFDNNYCRNPDGPRPWCYTLDPH 241
 Db 196 DYRGSDYRTESGKECQRMWDLQTPHAHPYKREKYPDKSLDNNYCRNPDSSSRPWCYTLDN 255
 Oy 242 TRNEYCAITCADNTMDTDVPLETTECTIOGSGEGYRGVNTIANGIPQCRMDSYQPHH 301
 Db 256 VEKEFRITKCKKORSNIEI--TSTCFKERGEYRGKANTTSGIPQCRMDSYQPHH 312
 Oy 302 DMTPEMFKCKDLAENYCRNPDSSSPWCTTIDNINVGCSOI-PNC-DMSHGDDCYRGNG 360
 Db 313 RLPEKYPCKGLDENYCRNPDSSSPWCTTIDNINVGCSOI-PNC-DMSHGDDCYRGNG 372

Oy 361 KNVGNLSQTRSGILTCSMWMDKNMEDLHRHIFWEPDASKLNNYCRNPDDAHGPWCYTGN 420
 Db 373 ELYGVSYSKTRKGIKCRMEKREKNDLESLD-QYLVPLENNYCRNPDROSHGWCYTMD 431
 Oy 421 PLIPMDYCPISRCGCD 436
 Db 432 PNTPEYCAIKPCEGE 447
 RESULT 12
 ID 091691 PRELIMINARY; PRT; 716 AA.
 AC 091691;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Growth factor Liverline.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ruiz I Alcaba A., Thery C.,
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U57455; AAB52574.1; -
 DR HSP; P00746; IFDP.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0003809; F:thrombin activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0007596; P:blood coagulation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR Pfam; PF00051; Kringleg; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; Trypsin; 1.
 DR PIRSF; PIRSF001152; HGF_MST1; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR Prodom; PD000395; Kringleg; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; TRYD_SPE; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS0070; KRINGLE_2; 4.
 DR PROSITE; PS00948; PAN; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR Hydroxylase; Kringleg; Protease; Serine protease.
 SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;
 Query Match 46.5%; Score 1198.5; DB 2; Length 716;
 Best Local Similarity 46.5%; Pred. No. 3.2e-82;
 Matches 205; Conservative 75; Mismatches 148; Indels 13; Gaps 6;
 Oy 5 RNTIEFFKSAKTTLIKIDPALKIKTKKYNVTADOCANRCTRNKGLPFTCKAFVFDKARQ 64
 Db 31 RALNDYQSKGLIELVHMNG-GVKOEIOSEIOVCAKQSD----LLDCRSFYVMKSGS 85
 Oy 65 CLMFPNMSGGVKKEFGHEFDLYENKDYIRNCITIGKRSYKTVSYTSGIKQCPWSSM 124
 Db 86 CHLLPMTQNSANVLLQRNVQYDLYQKDYIRDCVANGNTYRGVSKTSKGRCCWRMLK 145
 Oy 125 IPHEHY--RGKDLQENYCRNPRGEGGPGWCTSNPEVRYEVCDIPOCSEVECMTCNGE 181
 Db 146 FPHDHKFSPIHWPDLLENYCRNPDSPGPGWCTTIDKNIHQYCGIKKCEDAVCLTCNGE 205
 Oy 182 SYRGLMDHTESGKICQRMWDQTPHRRKFLPERYPDKGFDNNYCRNPDGPRPWCYTLDPH 241
 Db 206 DYRGSDYRTESGKECQRMWDLQTPHAHPYKREKYPDKSLDNNYCRNPDSSSRPWCYTLDN 265

```

QY 242 TRWEYCAIKTCADNTMNDTVPLETTBCTIOGQGEGRGVNTIWMGIPCGRWDSDQYPHH 301
DB 266 VEKEFRITCKCKQKRLSNIEI---TSTCFKRGEGRGKANTTSGICPCGRWDSQYPOSH 322
QY 302 DMTPENFKCDLRENYCRNPDGSESPWCFETTDNINIVGYSQIPNC-DMSHGDDCYRNG 360
DB 323 RLPEKRYPCGDLRENYCRNPDGSEAPWCFETTLPGMNAVCFQIKRCKDDVLEPDYCHNG 382
QY 361 KNYMGNLSTQSGLTCSMDKNMEDLHRHIFMEPPDASKLNENYCRNPDGDAHGPWCYTG 420
DB 383 ELYSGAVSKTRKIKCKRMEKNDLESLD-QPIVLEBENYCRNPDSDHGPWCYTTMD 441
QY 421 PLIPWDYCPISRCGDDTPTI 441
DB 442 PNTFPDYCAIKPCEGEKVLTL 462

RESULT 13
QY 090865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN Name=HGFL/MSP;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RX TISSUE=Liver;
RX MEDLINE=96029010; PubMed=7554499;
RT "Theory C., Sharpe M.J., Bailey S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGFL/MSP and c-met suggests new functions
RT during early chick development.";
RT Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: Belongs to peptidase family S1.
DB EMBL; X84043; CAA58862.1; -.
DB HSP; P00747; ICEA.
DR MEROPS; S01.977; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00089; Trypsin; 1.
DR PIRSF; PIRSF001152; HGF MST1; 1.
DR PRINTS; PR00723; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYD SP; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00948; PAN; 1.
DR PROSITE; PS02040; TRYPsin DOM; 1.
DR Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;

Query Match 45.9%; Score 1184; DB 2; Length 704;
Best Local Similarity 45.1%; Pred. No. 3.9e-81;
Matches 202; Conservative 72; Mismatches 150; Indels 24; Gaps 8;

QY 5 RNTIHFHKSAKTLIKIDPALKIKTKKVNTADQANCRTRNGLPFTCAVFDKARK 64
DB 20 RSLNDFQRLRGLELAAAPNEPPSPAPAHGAQCCQRCANRP---DCRAFHHEQSD 75
QY 65 CLWFPNMSGGVKEFGHEFDLYENKDYIRNCIIGKGRSYKGTVAITKGIKQCPWSSM 124

```

```

DB 76 COLLPMSQSPCARLQKNIHYDLYOKQFLRECIYANGTSYRGDTRTERGLRCHQWAT 135
QY 125 IPEHSY---RGKDIQENYCRNPDGSESPWCFETTDNINIVGYSQIPNC-DMSHGDDCYRNG 181
DB 136 TPDDHRLPRLNNGEENYCRNPDGSEAPWCFETTLPGMNAVCFQIKRCKDDVLEPDYCHNG 195
QY 182 SYRGMDHTESGKICQRMWDQTPHHKFLPERYPDKGFPDNYCRNPDGAPRPMCYTLDPH 241
DB 196 DYRGFVFDHESGTBECQRMWLDQHPHGRYPHDKYPEGLDDNYCRNPDGSEAPWCFETTLDPH 255
QY 242 TRWEYCAIKTCADNTMNDTVPLE-TTBTCTIOGQGEGRGVNTIWMGIPCGRWDSDQYPHH 300
DB 256 LREFEFCRIIVCKKRRP-----PINVTGCGYRGKGEGRGVNTIWMGIPCGRWDSDQYPHH 310
QY 301 DMTPENFKCDLRENYCRNPDGSESPWCFETTDNINIVGYSQIPNC-DMSHGDDCYRNG 359
DB 311 HHFVDSKTPCKDLOENYCRNPDGSEAPWCFETTLPGMNAVCFQIKRCKDDVLEPDYCHNG 370
QY 360 GKNYMGNLSTQSGLTCSMDKNMEDLHRHIFMEPPDASKLNENYCRNPDGDAHGP 414
DB 371 GERYHGVSKTRKIKCKRMEKNDLESLD-QPIVLEBENYCRNPDSDHGPWCYTTMD 425
QY 415 WCYTGNDPLIPWDYCPISRCGDDTPTI 442
DB 426 WCYTGNDPLIPWDYCPISRCGDDTPTI 453

RESULT 14
HGFL HUMAN
ID HGFL_HUMAN STANDARD; PRT; 711 AA.
AC P26927; Q13350; Q14870;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
DE stimulatory protein) (MSP) (Macrophage stimulating protein).
GN Name=MST1; Synonyms=HGFL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RX TISSUE=Liver;
RX MEDLINE=92002016; PubMed=1655021;
RT "Han S., Stuart L.A., Friesner Degen S.J.;
RT "Characterization of the DNFI52 locus on human chromosome 3:
RT identification of a gene coding for four kringle domains with homology
RT to hepatocyte growth factor.";
RT Biochemistry 30:9768-9780(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93340141; PubMed=8393443;
RT Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
RT "Cloning, sequencing, and expression of human macrophage stimulating
RT protein (MSP, MST1) confirms MSP as a member of the family of kringle
RT proteins and locates the MSP gene on chromosome 3.";
RT J. Biol. Chem. 268:15461-15468(1993).
RN [3]
RP SEQUENCE OF 230-247; 288-310; 326-341; 484-501; 530-549; 574-596 AND
RP 602-611, AND SUBUNIT.
RC TISSUE=Plasma;
RX Skeel A., Yoshimura T., Showalter S.D., Tanaka S., Appella E.,
RX Leonard E.J.;
RT "Macrophage stimulating protein: purification, partial amino acid
RT sequence, and cellular activity.";
RT J. Exp. Med. 173:1227-1234(1991).
CC -1- FUNCTION: Probably has no proteolytic activity, since crucial AA
CC characteristic of serine proteases catalytic sites are not
CC conserved.

```

CC -1- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
 CC disulfide bond.
 CC -1- PTM: May be cleaved after Arg-483, to yield two chains held
 CC together by disulfide bonds, or two separate polypeptides.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family. Plasminogen
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 kringle domains.
 CC -1- SIMILARITY: Contains 1 PAN domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, M74178; AAAS0165.1; -
 CC EMBL, U37055; AAC50471.1; -
 CC EMBL, L11924; AAAS9872.1; -
 CC PIR, A40331; A47136.
 CC HSP, P00766; 1CHG.
 CC MEROPS, S01.975; -
 CC Genew; HNC:7380; MST1.
 CC MIM, 142408; -
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan app.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR003966; Peptidase_S1A_pr.
 DR Pfam: PF00051; Kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; Trypsin; 1.
 DR PIRSF, PIRSF001152; HGF_MST1; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR PRINTS; PRO0018; KRINGLE.
 DR PRINTS; PRO1505; PROTHROMBIN.
 DR Prodom: PDO00395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00048; KRINGLE_2; 4.
 DR PROSITE; PS50948; PAN; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 KM Direct protein sequencing; Glycoprotein; Kringle; Polymorphism;
 KM Repeat; Serine protease homolog; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 711
 FT DOMAIN 21 105
 FT DOMAIN 110 186
 FT DOMAIN 191 268
 FT DOMAIN 283 361
 FT DOMAIN 370 448
 FT DOMAIN 484 711
 FT DISULFID 56 78
 FT DISULFID 60 66
 FT DISULFID 110 186
 FT DISULFID 131 169
 FT DISULFID 157 181
 FT DISULFID 194 324
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 283 361
 FT DISULFID 304 343
 FT DISULFID 332 355
 FT DISULFID 370 448
 FT DISULFID 391 431
 FT DISULFID 419 443
 FT DISULFID 468 588
 RT Interchain (By similarity).

FT DISULFID 507 523 By similarity.
 FT DISULFID 602 667 By similarity.
 FT DISULFID 632 646 By similarity.
 FT DISULFID 657 685 By similarity.
 FT CARBOHYD 72 72 N-linked [GlcNAc. . .] (Potential).
 FT CARBOHYD 296 296 N-linked [GlcNAc. . .] (Potential).
 FT CARBOHYD 615 615 N-linked [GlcNAc. . .] (Potential).
 FT VARIANT 13 13 Y -> C.
 FT VARIANT 13 13 /FTID=VAR_006531.
 FT VARIANT 212 212 C -> F.
 FT VARIANT 676 676 E -> K (in dbSNP:7798).
 FT VARIANT 676 676 /FTID=VAR_014569.
 FT CONFLICT 292 292 R -> G (in Ref. 3).
 FT CONFLICT 304 304 C -> E (in Ref. 3).
 FT CONFLICT 306 306 R -> E (in Ref. 3).
 FT CONFLICT 550 551 PS -> SL (in Ref. 3).
 FT CONFLICT 593 593 W -> E (in Ref. 3).
 FT CONFLICT 623 623 L -> F (in Ref. 2).
 SQ SEQUENCE 711 AA; 80379 MW; 596ED21F180290E4 CRC64;

Query Match 45.4%; Score 1171; DB 1; Length 711;
 Best Local Similarity 47.3%; Pred. No. 3.8e-80;
 Matches 196; Conservative 60; Mismatches 148; Indels 10; Gaps 4;

QY 33 VNTADQCANRCTRKNKGLPFTCKAFVFDKARQCLMFPNSNSGVKKEFGHEPDLYENKD 92
 DB 50 VADAECCAGRC-----GLMDGRAFHYNVSHGCCQLLPWTHGSPHTRLRSGRCDLFOKXD 105
 QY 93 YIRNCIIKKGKSYGKTSYTSYSGIKQCPWSSMIPIHESYR---GKLDQENCRNPRGREG 149
 DB 106 YVRICTIMNNGVGYGTATVATVGGJPCQAMSHKFPNDHYPTTLRNGLEENCRNPDGPG 165
 QY 150 GPWCPTSNPEVRYEVCDDPOCEVECMTCNGESRYGLMDHTESGKICRMWHOTPHRHKF 209
 DB 166 GPWCYTTDPAVRFGSCGKISCREAACVWCNEEYRGAVDRTESGECORMLQHPHQHPF 225
 QY 210 LPERYPDGFPDNYCRNPDGQPRPWCYTLDPHTREYCAIKTCADNTMNDTVPLEETEC 269
 DB 226 EPGKFLDQGLDDNYCRNPDGSERPWCYTTDQIREFCDLPRCSEAPROEA--TVSC 283
 QY 270 IQQGESEGRGVNTIMNGICPQWRDSDQYPRHHDMPENFKCDLRNCRNPDGSESPWC 329
 DB 284 FRKGEGYGRGTANTTAVPQQRWDAQIPHQHFRFPPEYACKDKRENCRNPDGSEAPWC 343
 QY 330 FTTDPNINVGYSQSPINC-DMSHGDCYRGNGKYNMGLSOTRSGLTCSMDKXNEDLAR 388
 DB 344 FTLRGMALAFCYQIRRCTDVRPDQCYHAGGEQYRGVSKTRKGVOCORSAETPRHKPQ 403
 QY 389 HIFWEPDASKLNNYCRNPDGDAHGPWCYTGNPILPMWDCPISRCGSDTPTTV 442
 DB 404 FTFTEPHAGQLEENCRNPDGDSHPWCYTMPDPTFPYCALRRCADQPPSIL 457
 RESULT 15
 ID P70006 PRELIMINARY; PRT; 717 AA.
 AC P70006;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 RX NCBI_TaxID=8355;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney, and Liver;
 RX MEDLINE=96404125; PubMed=8808403; DOI=10.1016/0925-4773(95)00458-0;
 RA Aberger F., Schmidt G., Richter K.;
 RT "The Xenopus genome of hepatocyte growth factor-like protein is

RT specifically expressed in the presumptive neural plate during
RT gastrulation.";
RL Mech. Dev. 54:23-37(1996) .
DR EMBL: Y08734; CAA69989.1; -.
DR HSSP: P00746; 1FDP.
DR MEROPS: S01.977; -.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0004263; F:chymotrypsin activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0003809; F:thrombin activity; IEA.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0007596; P:blood coagulation; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR011359; HGF_MST1.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; PAN_app.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR003966; Peptidase_S1A_pr.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR Pfam: PF00051; Kringle_4.
DR Pfam: PF00024; PAN_1.
DR Pfam: PF00089; Trypsin_1.
DR PIRSF: PIRSF001152; HGF_MST1_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR01505; PROTHROMBIN.
DR ProDom: PD000395; Kringle_4.
DR SMART: SM00130; KR_4.
DR SMART: SM00473; PAN_AP_1.
DR SMART: SM00020; TRYF_SPC_1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS00070; KRINGLE_2; 4.
DR PROSITE: PS02948; PAN_1.
DR PROSITE: PS02940; TRYPSIN_DOM; 1.
KW Hydrolyase; Kringle; Protease; Serine protease; Signal.
KW Hydrolyase; Kringle; Protease; Signal.
FT SIGNAL 1 29 potential.
FT CHAIN 1 29 hepatocyte growth factor-like protein.
SQ SEQUENCE 717 AA; 82017 MW; 6F877AA32C6CDD54 CRC64;

Query Match 45.1%; Score 1163.5; DB 2; Length 717;
Best Local Similarity 45.1%; Pred. No. 1.4e-79;
Matches 199; Conservative 77; Mismatches 152; Indels 13; Gaps 6;

QY 5 RNTIHFEXKSAKTLIKIPALKIKTKKXNTADQCANRCRNKGLPFTCKAFVFDKARQ 64
DB 32 KSLANDYQSKGIELVHMBG-GVKQEVQAEIQICAKQSD---LLDCRSFDPYMKSSQS 86
QY 65 CLWFPNMSNGVYKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVAITKSGIKQPSMSM 124
DB 87 CRLLPWTQNSNVLQRNQYDLYQKKDYIRDCVGNNGMYRTGTVSKTKNGRTQCHWRLLK 146
QY 125 IPHEHSY---RGKDLOENTCRNPRGEGGFWCTSNPEVRYEYCDIPQCEVECMTCNGE 181
DB 147 FPHDHKFSPTHWLEENYCRNPDSDEGLMCYTTDKNIRHOYCGIKCEDAVCLTCNGE 206
QY 182 SYRGGLDHTESGKICORWDHQPFRHKPLPERYPDKGPDNDYCRNPDSQPRPWCYTLDPH 241
DB 207 DYRGSDYRTESGECQRMDOABPHYPERKYPDKSLDNYCRNPDSSERPWCYTTDPN 266
QY 242 TRWEYCAIKTCADNTNADTVPLETTECIQGGEGYRGVTNITWNGIPCCORMDSQYRHH 301
DB 267 VEREFCHITKCKKQRIISNIEI--TSTCKREGEGRKANNTTSGIPCCORMDTQAPVH 323
QY 302 DMTPEPFCKDLRENTCRNPDSGSPWCTTDPNIRVYCSQIPNC-DMSHGDCYRANG 360
DB 324 RFLPERYPCGKIDENYCRNPVGSSEAPCTTLKNMRMAYCFQIKRCTDDVVEPECYHNG 383
QY 361 KVMYGNLSQTRSGLTCSMDKNMEDLRHIFWEPDASKLNENYCRNPDDAHGPPWCYGN 420
DB 384 ELYRGVSKTRKGIKCRWEKONDLELSLA-OPYLVLEENYCRNPDRDSHGPPWCYTN 442

QY 421 PLIPWDYCPISRCEGDTTPTI 441
DB 443 PNTPEFCALKPKCAGDKVLTL 463
Search completed: October 3, 2005, 06:08:09
Job time : 64.6344 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:29 ; Search time 170.535 seconds
(without alignments)
1002.420 Million cell updates/sec

Title: US-09-674-377B-2

Perfect score: 2579
Sequence: 1 EKRKNITHEFKSAKTLTI.....IPWDYCPISRCRGTPTTIV 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20015:*
5: geneseqp20025:*
6: geneseqp20035:*
7: geneseqp20035:*
8: geneseqp20045:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2576	99.9	442	7	ADM83267 Human NK4
2	2576	99.9	723	2	AAR07144 Tumour cy
3	2576	99.9	723	2	AAR15624 Human leu
4	2576	99.9	723	2	AAR25677 Recombina
5	2576	99.9	723	2	AAR57027 Human mod
6	2576	99.9	723	2	AAR57028 Human mod
7	2576	99.9	723	2	AAR57026 Human wll
8	2576	99.9	723	2	AAR82685 Tumour cy
9	2576	99.9	723	2	AAR76690 Human pla
10	2576	99.9	723	2	AAY06621 HGF-MSP h
11	2576	99.9	723	2	AAY06622 HGF-MSP h
12	2576	99.9	1109	7	ADDS9672 Chimeric
13	2574	99.8	442	3	AAY57174 N-termina
14	2573	99.8	723	2	AAR21142 Human TCF
15	2566	99.5	723	2	AAR59923 Human leu
16	2565	99.5	504	2	AAR32710 Haematopo
17	2563.5	99.4	447	7	ADBS37393 Human hep
18	2563.5	99.4	447	7	ADM83266 Human NK4
19	2563.5	99.4	697	3	AAY8485 Hepatocyt
20	2563.5	99.4	697	3	AAY59030 Sequence
21	2563.5	99.4	697	4	AAB45838 Nucleic a
22	2563.5	99.4	697	4	AAB45838 Nucleic a
23	2563.5	99.4	728	2	AAR25676 Recombina
24	2563.5	99.4	728	2	AAR20005 Human hep
25	2563.5	99.4	728	2	AAR40862 Competati

26	2563.5	99.4	728	2	AAR40863 Competati
27	2563.5	99.4	728	2	AAR42062 Vascular
28	2563.5	99.4	728	2	AAR87522 Mutant he
29	2563.5	99.4	728	2	AAR87524 Mutant he
30	2563.5	99.4	728	2	AAR87525 Mutant he
31	2563.5	99.4	728	2	AAR87523 Mutant he
32	2563.5	99.4	728	2	AAR00340 Wild type
33	2563.5	99.4	728	2	AAR00338 Human hep
34	2563.5	99.4	728	2	AAR59922 Human leu
35	2563.5	99.4	728	2	AAR58696 Human hep
36	2563.5	99.4	728	2	AAR42998 Recombina
37	2563.5	99.4	728	2	AAR39207 Human hep
38	2563.5	99.4	728	6	ABP58128 Human hep
39	2563.5	99.4	728	7	ADB61534 Hepatocyt
40	2563.5	99.4	728	8	ADM97652 Human hep
41	2563.5	99.4	728	8	ADR89322 Human hep
42	2563.5	99.4	951	8	ADS19044 Chimeric
43	2561.5	99.3	447	3	AAY57173 N-termina
44	2557.5	99.2	727	2	AAR10656 Hepatic p
45	2557.5	99.2	728	2	AAR15623 Human leu

ALIGNMENTS

RESULT 1	ADM83267	
ID	ADM83267	standard; protein; 442 AA.
XX	AC	ADM83267;
XX	AC	ADM83267;
DT	XX	18-NOV-2004 (first entry)
XX	XX	Human NK4 protein #2.
DE	XX	
XX	KW	NK4 gene; neovascularisation; ovarian cancer; pancreatic cancer;
XX	KW	stomach cancer; gallbladder cancer; renal cancer; prostatic cancer;
XX	KW	breast cancer; esophageal cancer; hepatic cancer; oral cavity cancer;
XX	KW	colon cancer; colorectal cancer; sarcoma; glioma; melanoma; gene therapy;
XX	KW	human.
OS	XX	Homo sapiens.
XX	PN	US2003162736-A1.
XX	PN	26-AUG-2003.
PD	XX	27-SEP-2002; 2002US-00255649.
PF	XX	25-FEB-2002; 2002JP-0048644.
PR	XX	(NARA/) NAKAMURA T.
PA	XX	(MATS/) MATSUMOTO K.
PI	XX	Nakamura T, Matsumoto K;
DR	XX	WPI: 2003-897946/82.
DR	XX	N-PSDB; ADM83265.
PT	XX	New NK4 gene, useful for producing a medicament for preventing or
PT	XX	treating diseases caused by neovascularization, e.g. cancer.
PS	XX	Disclosure; Page 17-19; 27pp; English.
CC	XX	The present invention provides a therapeutic agent NK4 polypeptides and
CC	XX	their encoding polynucleotides. The invention is useful for producing a
CC	XX	medicament for preventing or treating diseases caused by
CC	XX	neovascularisation such as ovarian cancer, pancreatic cancer, stomach
CC	XX	cancer, gallbladder cancer, renal cancer, prostatic cancer, breast
CC	XX	cancer, esophageal cancer, hepatic cancer, oral cavity cancer, colon
CC	XX	cancer, colorectal cancer, sarcoma, glioma and melanoma. The invention is
CC	XX	also useful in gene therapy. The present sequence is human NK4 protein.

SQ Sequence 442 AA;

Query Match 99.9%; Score 2576; DB 7; Length 442;
 Best Local Similarity 99.8%; Pred. No. 8.5e-166;
 Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPFTCAAFVFDK 60
 DB 1 QKRRNTIHEFKSAKTTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPFTCAAFVFDK 60
 QY 61 ARKOCILMPFNSMSGVKKERGFHEFDLYENKDYINNCIIIGKRSYKGVSTITKGIKOP 120
 DB 61 ARKOCILMPFNSMSGVKKERGFHEFDLYENKDYINNCIIIGKRSYKGVSTITKGIKOP 120
 QY 121 WSSMI PHEHSYRGKDLQENYCRNPRGEGGFWCTSNPEVREYECDI PQCEVECMTCNG 180
 DB 121 WSSMI PHEHSYRGKDLQENYCRNPRGEGGFWCTSNPEVREYECDI PQCEVECMTCNG 180
 QY 181 ESYRGLMDHTESGKICQRMHQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 240
 DB 181 ESYRGLMDHTESGKICQRMHQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 240
 QY 241 HTRWEYCAIKTCADNTMDTVPLETTTCICGQGGEGYGTVNTIWNGLPCQRMBSQYFHE 300
 DB 241 HTRWEYCAIKTCADNTMDTVPLETTTCICGQGGEGYGTVNTIWNGLPCQRMBSQYFHE 300
 QY 301 HDMPENFKCKDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCDSHGDQCYRGNG 360
 DB 301 HDMPENFKCKDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCDSHGDQCYRGNG 360
 QY 361 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 420
 DB 361 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 420
 QY 421 PLIPWDYCPISRCBGDTTPTIV 442
 DB 421 PLIPWDYCPISRCBGDTTPTIV 442

RESULT 2

AAR07144 standard; protein; 723 AA.

XX AAR07144;
 AC AAR07144;
 XX 25-MAR-2003 (revised)
 DT 31-JAN-1991 (first entry)
 XX Tumour cytotoxic factor II.
 DE
 XX TCF II; glycoprotein; fibroblasts; anticancer agent.
 KW
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..29
 XX
 XX WO9010651-A.
 XX 20-SEP-1990.
 PD
 XX 10-MAR-1989; 89JP-00058631.
 PF
 XX 10-MAR-1989; 89JP-00058631.
 PR 16-JAN-1990; 90JP-00006692.
 XX
 XX (SNOW) SNOW BRAND MILK PROD CO LTD.
 PA
 XX Higashio K, Mitsuda S, Shima N, Itagaki Y, Nagao M,
 PI WPI; 1990-304990/40.
 DR P-PSDB; AAR07144.
 XX

PT Human glyco:protein having antitumour activity - for inducing
 PT differentiation of leukaemia cells and enhancing cell mediated immunity.
 XX
 PS Disclosure; Fig 15; 73pp; Japanese.

CC The sequence was deduced from the cDNA obt'd. from a clone isolated from a
 CC cDNA library prepd. from RNA ext'd. from fibroblast IMR-90 cells (ATCC CCL
 CC -186). The N-terminal of the alpha chain is unclear. The sequence shows
 CC homology to hGF except between AAs 162 and 166 of this sequence which is
 CC absent from the TCF-II sequence. The DNA can be used to produce the TCF
 CC II by recombinant DNA technology. TCF II is an anticancer agent and
 CC induces differentiation of leukaemia cells, enhances cell-mediated
 CC immunity, and accelerates the proliferation of human blood vessel
 CC endothelial cells and hepatic parenchymal cells. (updated on 25-MAR-2003
 CC to correct PR field.) (updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 723 AA;

Query Match 99.9%; Score 2576; DB 2; Length 723;
 Best Local Similarity 99.8%; Pred. No. 1.4e-163;
 Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPFTCAAFVFDK 60
 DB 32 QKRRNTIHEFKSAKTTLLIKIDPALKIKTKKVTADOCANRCTRNKGLPFTCAAFVFDK 91
 QY 61 ARKOCILMPFNSMSGVKKERGFHEFDLYENKDYINNCIIIGKRSYKGVSTITKGIKOP 120
 DB 92 ARKOCILMPFNSMSGVKKERGFHEFDLYENKDYINNCIIIGKRSYKGVSTITKGIKOP 151
 QY 121 WSSMI PHEHSYRGKDLQENYCRNPRGEGGFWCTSNPEVREYECDI PQCEVECMTCNG 180
 DB 152 WSSMI PHEHSYRGKDLQENYCRNPRGEGGFWCTSNPEVREYECDI PQCEVECMTCNG 211
 QY 181 ESYRGLMDHTESGKICQRMHQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 240
 DB 212 ESYRGLMDHTESGKICQRMHQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 271
 QY 241 HTRWEYCAIKTCADNTMDTVPLETTTCICGQGGEGYGTVNTIWNGLPCQRMBSQYFHE 300
 DB 272 HTRWEYCAIKTCADNTMDTVPLETTTCICGQGGEGYGTVNTIWNGLPCQRMBSQYFHE 331
 QY 301 HDMPENFKCKDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCDSHGDQCYRGNG 360
 DB 332 HDMPENFKCKDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCDSHGDQCYRGNG 391
 QY 361 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 420
 DB 392 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 451
 QY 421 PLIPWDYCPISRCBGDTTPTIV 442
 DB 452 PLIPWDYCPISRCBGDTTPTIV 473

RESULT 3

AAR15624 standard; protein; 723 AA.

XX AAR15624;
 AC AAR15624;
 XX 25-MAR-2003 (revised)
 DT 18-MAR-1992 (first entry)
 XX Human leukocyte-derived HGF encoded by clone HLC3.
 DE Hepatocyte growth factor; liver; hepatoma.
 KW
 XX Homo sapiens.
 OS
 XX EP461560-A.
 XX

PD 18-DEC-1991.
XX
PF 07-JUN-1991; 91EP-00109369.
XX
PR 11-JUN-1990; 90JP-00152474.
XX
PA (TOYM) TOYO BOSEKI KK.
XX
PI Nakamura T, Hagiya M, Seki T, Shimomishi M, Shimizu S, Ihara I;
PI Sakauchi M, Asami O;
XX
DR WPI: 1991-370578/51.
DR N-PSDB; AAQ15177.
PT Recombinant human leukocyte-derived hepatocyte growth factor - with DNA
PT encoding it, recombinant expression vectors and transformant cells
PT expressing it.
XX
PS Claim 3; Fig 3; 33pp; English.
XX
CC The sequence was deduced from a portion of HUC3, one of two clones, (for
CC HUC2 see AAR15623) isolated from a cDNA library prepd. from mRNA
CC extracted from human leukocytes. HUC2 has almost the same sequence as
CC HUC3 except for five residues (162-166) in HUC2 which do not appear in
CC HUC3. HUC3 shows similar characteristics to the human liver-derived HGF
CC identified in Nature, 342, 440, 1989, but differs at 14 positions in the
CC amino acid sequence. The DNA sequence can be expressed and the resulting
CC protein, recombinant HGF, used in hepa- tocyte cultivation, liver
CC regeneration, hepatocyte research, esp. into the mechanism of hepatoma,
CC and to prepare anti-HGF antibodies for diagnosis and therapy. (Updated on
CC 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 723 AA;
XX
Query Match 99.9%; Score 2576; DB 2; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.4e-163;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRNTTIEFKSAKTLTIKIDPALKITKVNNTADOCANRCTRNKGIPFTCKAFVFDK 60
DB 32 QRKRNTTIEFKSAKTLTIKIDPALKITKVNNTADOCANRCTRNKGIPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTIVSITKSGIKCOP 120
DB 92 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTIVSITKSGIKCOP 151
QY 121 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVDIPOCSVEECMTONG 180
DB 152 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVDIPOCSVEECMTONG 211
QY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHAKFLPERYPDKGFDDNYCRNPDGQRPWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIRHAKFLPERYPDKGFDDNYCRNPDGQRPWCYTLDP 271
QY 241 HTRWEYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGTVNTIWNGLPCQRMDSQYPHE 300
DB 272 HTRWEYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGTVNTIWNGLPCQRMDSQYPHE 331
QY 301 HDMTPENFKCDLRNRYCRNPDGSSPWCFTTDPNIRVGYCSQIPNCMSHGQDCYRNG 360
DB 332 HDMTPENFKCDLRNRYCRNPDGSSPWCFTTDPNIRVGYCSQIPNCMSHGQDCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPWCYTGN 451
QY 421 PLIPWDYCPISRCBGDTTPTIV 442
DB 452 PLIPWDYCPISRCBGDTTPTIV 473
RESULT 4
AAR25677

ID AAR25677 standard; protein; 723 AA.
XX
AC AAR25677;
XX
DT 20-JUN-1993 (first entry)
XX
DE Recombinant human hepatocyte growth factor.
XX
KW HGF; enhance growth; preparing transgenic animals; hepatic disease;
KW clinical diagnostic reagent; drug.
XX
OS Homo sapiens.
XX
PN JP04183394-A.
XX
PD 30-JUN-1992.
XX
PF 19-NOV-1990; 90JP-00314548.
XX
PR 19-NOV-1990; 90JP-00314548.
XX
PA (TOYM) TOYOBO KK.
PA (NAKA/) NAKAMURA T.
XX
DR WPI: 1992-265591/32.
XX
PT Recombinant human hepatocyte growth factor and DNA encoding it - useful
PT for diagnosis and treatment of hepatic disease and transgenic animal
XX
PS Disclosure; Page 13; 28pp; Japanese.
XX
CC This sequence represents a recombinant human hepatocyte growth factor. It
CC has physiological activity, and using it enhanced growth of hepatocytes
CC is possible. It is useful as a clinical diagnostic reagent, or a drug for
CC treating hepatic disease. See also AAR25676-92, AAQ26713-27
XX
SQ Sequence 723 AA;
XX
Query Match 99.9%; Score 2576; DB 2; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.4e-163;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRNTTIEFKSAKTLTIKIDPALKITKVNNTADOCANRCTRNKGIPFTCKAFVFDK 60
DB 32 QRKRNTTIEFKSAKTLTIKIDPALKITKVNNTADOCANRCTRNKGIPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTIVSITKSGIKCOP 120
DB 92 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTIVSITKSGIKCOP 151
QY 121 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVDIPOCSVEECMTONG 180
DB 152 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVDIPOCSVEECMTONG 211
QY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHAKFLPERYPDKGFDDNYCRNPDGQRPWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIRHAKFLPERYPDKGFDDNYCRNPDGQRPWCYTLDP 271
QY 241 HTRWEYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGTVNTIWNGLPCQRMDSQYPHE 300
DB 272 HTRWEYCAIKTCADNTMNDTDVPLETTECTIOGQEGYRGTVNTIWNGLPCQRMDSQYPHE 331
QY 301 HDMTPENFKCDLRNRYCRNPDGSSPWCFTTDPNIRVGYCSQIPNCMSHGQDCYRNG 360
DB 332 HDMTPENFKCDLRNRYCRNPDGSSPWCFTTDPNIRVGYCSQIPNCMSHGQDCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPWCYTGN 451
QY 421 PLIPWDYCPISRCBGDTTPTIV 442

Db 452 PLIPWDYCPISRCBGDTTPTIV 473

RESULT 5
AAR57027
ID AAR57027 standard; protein; 723 AA.
XX
AC AAR57027;
XX
DT 25-MAR-2003 (revised)
DT 19-MAR-1995 (first entry)
XX
DE Human modified tumor cytotoxic factor.
XX
KW Tumor cytotoxic factor; hepatocyte growth factor; antitumor.
XX
OS Homo sapiens.
XX
PN WO9414845-A1.
XX
PD 07-JUL-1994.
XX
PF 27-DEC-1993; 93WO-JP001905.
XX
PR 28-DEC-1992; 92JP-00359747.
XX
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Yamaguchi K, Shima N, Murakami A, Goto M, Tsuda E, Masunaga H;
PI Takahira R, Oogaki F, Ueda M, Higashio K;
XX
DR WPI; 1994-234619/28.
XX
PT Modified tumour cytotoxic factors with loss of binding of N-linked
PT oligosaccharide chain - have longer biological half-life without loss of
PT biological activity.
XX
PS Claim 5; Page 33; 59pp; English.
XX
CC This modified TCF sequence has a prolonged biological half-life without
CC any loss of biological activity, is obtained by altering the amino acid
CC residue at position 563 from Ser in the wild-type to Ala. (Updated on 25-
CC MAR-2003 to correct PN field.)
XX
SQ Sequence 723 AA;

Query Match 99.9%; Score 2576; DB 2; Length 723;
Best Local Similarity 99.8%; Pred. No. 1,4e-163;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTTHERFKSAKTTLLIKIDPALKIKTKKVNTPADOCANRCTRNKGLPTCKAFVFDK 60
:|||||
DB 32 QKRKRTTHERFKSAKTTLLIKIDPALKIKTKKVNTPADOCANRCTRNKGLPTCKAFVFDK 91

QY 61 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 120
:|||||
DB 92 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 151

QY 121 WSSMTIPHESYRGKDLOENYCRNPRGEGGFCFTSNPEVRYEVCIDIPOCSEVECMTCNG 180
:|||||
DB 152 WSSMTIPHESYRGKDLOENYCRNPRGEGGFCFTSNPEVRYEVCIDIPOCSEVECMTCNG 211

QY 181 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 240
:|||||
DB 212 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 271

QY 241 HTRWEYCAIKTCADNTMDTDPVLETTBECIOGQSEGYRGTVNTIWMGIPQCRWDSQYRHE 300
:|||||
DB 272 HTRWEYCAIKTCADNTMDTDPVLETTBECIOGQSEGYRGTVNTIWMGIPQCRWDSQYRHE 331

QY 301 HMTPTNFYCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 360
:|||||
DB 332 HMTPTNFYCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 391

QY 361 KNYMGNLSQTRSSGLTCSMWDKNMEDLHRHIFWEPDASKLNNYCRNPDDBAHGFWCYTGN 420
:|||||
DB 392 KNYMGNLSQTRSSGLTCSMWDKNMEDLHRHIFWEPDASKLNNYCRNPDDBAHGFWCYTGN 451

QY 421 PLIPWDYCPISRCBGDTTPTIV 442
:|||||
DB 452 PLIPWDYCPISRCBGDTTPTIV 473

RESULT 6
AAR57028
ID AAR57028 standard; protein; 723 AA.
XX
AC AAR57028;
XX
DT 25-MAR-2003 (revised)
DT 19-MAR-1995 (first entry)
XX
DE Human modified tumor cytotoxic factor.
XX
KW Tumor cytotoxic factor; hepatocyte growth factor; antitumor.
XX
OS Homo sapiens.
XX
PN WO9414845-A1.
XX
PD 07-JUL-1994.
XX
PF 27-DEC-1993; 93WO-JP001905.
XX
PR 28-DEC-1992; 92JP-00359747.
XX
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Yamaguchi K, Shima N, Murakami A, Goto M, Tsuda E, Masunaga H;
PI Takahira R, Oogaki F, Ueda M, Higashio K;
XX
DR WPI; 1994-234619/28.
XX
PT Modified tumour cytotoxic factors with loss of binding of N-linked
PT oligosaccharide chain - have longer biological half-life without loss of
PT biological activity.
XX
PS Claim 6; Page 33; 59pp; English.
XX
CC This modified TCF sequence has a prolonged biological half-life without
CC any loss of biological activity, is obtained by altering the amino acid
CC residue at position 563 from Ser in the wild-type to Ala. (Updated on 25-
CC MAR-2003 to correct PN field.)
XX
SQ Sequence 723 AA;

Query Match 99.9%; Score 2576; DB 2; Length 723;
Best Local Similarity 99.8%; Pred. No. 1,4e-163;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTTHERFKSAKTTLLIKIDPALKIKTKKVNTPADOCANRCTRNKGLPTCKAFVFDK 60
:|||||
DB 32 QKRKRTTHERFKSAKTTLLIKIDPALKIKTKKVNTPADOCANRCTRNKGLPTCKAFVFDK 91

QY 61 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 120
:|||||
DB 92 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 151

QY 121 WSSMTIPHESYRGKDLOENYCRNPRGEGGFCFTSNPEVRYEVCIDIPOCSEVECMTCNG 180
:|||||
DB 152 WSSMTIPHESYRGKDLOENYCRNPRGEGGFCFTSNPEVRYEVCIDIPOCSEVECMTCNG 211

QY 181 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 240
:|||||
DB 212 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 271

```
QY 241 HTRMEYCAIKTCADNTMNDTVPLETTBCTIOGQSGYRGTVNTWNGIPCQWDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTBCTIOGQSGYRGTVNTWNGIPCQWDSQYPHE 331
QY 301 HMTPEPFCKDRLRENYCRNPDGSSPMCFITDPNIRVGCQI PNCDSHODCYRNG 360
DB 332 HMTPEPFCKDRLRENYCRNPDGSSPMCFITDPNIRVGCQI PNCDSHODCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 451
QY 421 PLIPMDYCPISRCEGDTPTTIV 442
DB 452 PLIPMDYCPISRCEGDTPTTIV 473

RESULT 7
AAR57026
ID AAR57026 standard; protein; 723 AA.
XX
AC AAR57026;
XX
DT 25-MAR-2003 (revised)
DT 19-MAR-1995 (first entry)
XX
DE Human wild-type tumor cytotoxic factor.
XX
KM Tumor cytotoxic factor; hepatocyte growth factor; antitumor.
XX
OS Homo sapiens.
XX
PN WO9414845-A1.
XX
PD 07-JUL-1994.
XX
PF 27-DEC-1993; 93WO-JP001905.
XX
PR 28-DEC-1992; 92JP-00359747.
XX
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
XX
PI Yamaguchi K, Shima N, Murakami A, Goto M, Tsuda E, Masunaga H;
PI Takahira R, Oogaki F, Ueda M, Higashio K;
XX
DR WPI; 1994-234619/28.
XX
PT Modified tumour cytotoxic factors with loss of binding of N-linked
PT oligosaccharide chain - have longer biological half-life without loss of
PT biological activity.
XX
PS Disclosure; Page 29; 59pp; English.
XX
CC This TCF sequence may be modified in that the AA residues responsible for
CC binding of N-linked oligosaccharide chains are prevented so that at least
CC one of the N-linked oligosaccharide chains is prevented from binding to
CC TCF. Ser-563 may be substituted with Ala (AAR57027) or Asp-289
CC substituted with Gln (AAR57028). (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 723 AA;

Query Match 99.9%; Score 2576; DB 2; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.4e-163;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRRNTIEFKKSAKTLIKIDPALKIKTKKVNADQCANRTRNGLPFTCKAFVFDK 60
DB 32 QKRRNTIEFKKSAKTLIKIDPALKIKTKKVNADQCANRTRNGLPFTCKAFVFDK 91
QY 61 ARKQCIWPFPPNSGSSGVKKEFGHEPDLLENKYIRNCITIGKRSYKGTYSITKSGIKCOP 120
DB 92 ARKQCIWPFPPNSGSSGVKKEFGHEPDLLENKYIRNCITIGKRSYKGTYSITKSGIKCOP 151
```

```
QY 121 WSSMIPHEHSYKGLQENYCRNPRGEGGPMCFITSNPEVRYEVCIDIPOCSEVECTCNG 180
DB 152 WSSMIPHEHSYKGLQENYCRNPRGEGGPMCFITSNPEVRYEVCIDIPOCSEVECTCNG 211
QY 181 ESYRGLMDHTESGKICQWMDHOTPHRHKFLPERYRDKGFDNDYCRNPDQPPRWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQWMDHOTPHRHKFLPERYRDKGFDNDYCRNPDQPPRWCYTLDP 271
QY 241 HTRMEYCAIKTCADNTMNDTVPLETTBCTIOGQSGYRGTVNTWNGIPCQWDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTBCTIOGQSGYRGTVNTWNGIPCQWDSQYPHE 331
QY 301 HMTPEPFCKDRLRENYCRNPDGSSPMCFITDPNIRVGCQI PNCDSHODCYRNG 360
DB 332 HMTPEPFCKDRLRENYCRNPDGSSPMCFITDPNIRVGCQI PNCDSHODCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCYTGN 451
QY 421 PLIPMDYCPISRCEGDTPTTIV 442
DB 452 PLIPMDYCPISRCEGDTPTTIV 473

RESULT 8
AAR82685
ID AAR82685 standard; protein; 723 AA.
XX
AC AAR82685;
XX
DT 24-APR-1996 (first entry)
XX
DE Tumour cytotoxic factor-II.
XX
KM Tumour cytotoxic factor-II; TCF-II; hepatocyte growth factor; HGF;
KM monoclonal antibody; MAD; hybridoma; antitumour agent.
XX
OS Homo sapiens.
XX
PN EP672685-A2.
XX
PD 20-SEP-1995.
XX
PF 17-MAR-1995; 95EP-00301800.
XX
PR 18-MAR-1994; 94JP-00074263.
XX
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
XX
PI Higashio K, Shima N, Oogaki F;
XX
DR WPI; 1995-329569/43.
XX
PT Monoclonal antibody to human tumour cytotoxic factor-II - has no affinity
PT to human hepatocyte growth factor, is used for specific determ. or
PT purification of TCF-II.
XX
PS Disclosure; Page 14-17; 19pp; English.
XX
CC This sequence represents human tumour cytotoxic factor-II (TCF-II). TCF-
CC II has potent antitumour activity and growth stimulation activity. TCF-II
CC is a member of the hepatocyte growth factor family (HGF). This sequence
CC is used as an antigen to produce TCF-II specific antibodies that show no
CC affinity to human HGF. The monoclonal antibody (MAB) was produced using
CC hybridoma techniques. The MABs were screened using solid phase ELISA with
CC an increased selectivity. The MAB can be used for the selective
CC determination or purification of TCF-II without any influence from the
CC presence of HGF. The TCF-II can be used as an antitumour agent, a wound
CC healing agent and as a marker for the diagnosis of diseases
XX
SQ Sequence 723 AA;
```

Query Match 99.9%; Score 2576; DB 2; Length 723;
Best Local Similarity 99.8%; Pred. No. 1,4e-163;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEPKKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCAFAVPDK 60
DB :
DB 32 ORKRRNTIHEPKKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCAFAVPDK 91
QY 61 ARKQCLMPFPMSSSGVKKERGFEDLYENKDYIRNCLIGKRSYKGVSTTKSGIKQOP 120
DB 92 ARKQCLMPFPMSSSGVKKERGFEDLYENKDYIRNCLIGKRSYKGVSTTKSGIKQOP 151
QY 121 WSSMIPIHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVREYECDIPOCSEVECMTCNG 180
DB 152 WSSMIPIHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVREYECDIPOCSEVECMTCNG 211
QY 181 ESYRGLMDHTESGKICQRMHDQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMHDQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 271
QY 241 HTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGTVNTIWNGLPCQRMDSQYFHE 300
DB 272 HTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGTVNTIWNGLPCQRMDSQYFHE 331
QY 301 HDMTPENFKCDLRENYCRNPDGSESPPCFTTDNIRVGYCSQIPNCDMSHGQDCYRNG 360
DB 332 HDMTPENFKCDLRENYCRNPDGSESPPCFTTDNIRVGYCSQIPNCDMSHGQDCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKXMDLHRHIFWEPDASKLKNENYCRNPDDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGLTCSMWDKXMDLHRHIFWEPDASKLKNENYCRNPDDAHGPWCYTGN 451
QY 421 PLIPWDYCPISRCGDDTTPTIV 442
DB 452 PLIPWDYCPISRCGDDTTPTIV 473

RESULT 9
AAW76690
ID AAW76690 standard; protein; 723 AA.
XX AAW76690;
AC
XX
DT 25-MAR-2003 (revised)
DT 29-JAN-1999 (first entry)
XX
DE Human plasminogen-like growth factor.
XX
KW Human; plasminogen-like growth factor; PLGF; endothelial cell;
KW epithelial cell; splice variant; wound healing.
XX
OS Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 175
FT /note= "Encoded by ACT"

US5621223-A.
XX
PD 13-OCT-1998.
XX
PF 05-MAY-1994; 94US-00238742.
XX
PR 14-SEP-1990; 90US-00582063.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Aaronson SA, Rubin JS, Chan AM;
XX
DR WPI, 1998-567711/48.
XX
DR N-PSDB; AAV61952.
XX

PT Stimulation of endothelial or epithelial cell growth - with plasminogen-
PT like growth factor protein.
XX
PS Claim 3; Fig 4C; 19pp; English.
XX
CC The Human plasminogen-like growth factor PLGF is used in the stimulation
CC of the growth of endothelial cells or epithelial cells other than
CC hepatocytes. This involves contacting the cells with the PLGF protein or
CC its allelic or splice variants. The growth factor is useful for
CC stimulating wound healing in endothelial or epithelial tissue. (Updated
CC on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 723 AA;

Query Match 99.9%; Score 2576; DB 2; Length 723;
Best Local Similarity 99.8%; Pred. No. 1,4e-163;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEPKKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCAFAVPDK 60
DB :
DB 32 ORKRRNTIHEPKKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCAFAVPDK 91
QY 61 ARKQCLMPFPMSSSGVKKERGFEDLYENKDYIRNCLIGKRSYKGVSTTKSGIKQOP 120
DB 92 ARKQCLMPFPMSSSGVKKERGFEDLYENKDYIRNCLIGKRSYKGVSTTKSGIKQOP 151
QY 121 WSSMIPIHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVREYECDIPOCSEVECMTCNG 180
DB 152 WSSMIPIHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVREYECDIPOCSEVECMTCNG 211
QY 181 ESYRGLMDHTESGKICQRMHDQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMHDQTPHRAKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDP 271
QY 241 HTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGTVNTIWNGLPCQRMDSQYFHE 300
DB 272 HTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGTVNTIWNGLPCQRMDSQYFHE 331
QY 301 HDMTPENFKCDLRENYCRNPDGSESPPCFTTDNIRVGYCSQIPNCDMSHGQDCYRNG 360
DB 332 HDMTPENFKCDLRENYCRNPDGSESPPCFTTDNIRVGYCSQIPNCDMSHGQDCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKXMDLHRHIFWEPDASKLKNENYCRNPDDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGLTCSMWDKXMDLHRHIFWEPDASKLKNENYCRNPDDAHGPWCYTGN 451
QY 421 PLIPWDYCPISRCGDDTTPTIV 442
DB 452 PLIPWDYCPISRCGDDTTPTIV 473

RESULT 10
AAV06621
ID AAV06621 standard; protein; 729 AA.
XX AAV06621;
AC
XX
DT 26-OCT-1999 (first entry)
DT
XX
DE HGF-MSP hybrid protein alphabet-1 factor.
XX
KW Alphabet-1; human; hepatocyte growth factor; HGF; MSP;
KW macrophage stimulating protein; apoptosis; chemotherapy; toxicity;
KW therapy.
XX
OS Homo sapiens.
OS
XX
FH Homo Sapiens.
FH
FT Domain
FT Key Location/Qualifiers
FT Peptide 1..487
FT /note= "HGF alpha chain"
FT 1..31
FT /note= "HSP alpha signal peptide"

FT	Protein	32..729
FT	/note="mature protein"	
FT	Cleavage-site	488..491
FT	/note="HGF cleavage site"	
FT	Domain	492..717
FT	/note="MSP beta chain"	
FT	Peptide	718..729
FT	/note="polyhistidine tag"	
XX		
PN	WO938968-A1.	
XX		
PD	05-AUG-1999.	
XX		
PP	28-JAN-1999;	99WO-EP000502.
XX		
PR	30-JAN-1998;	98IT-M1000180.
XX		
PA	(DOMP-) DOMPE SPA.	
XX		
PI	Medico E, Michieli P, Collesi C, Caselli G, Comoglio P;	
XX		
DR	WPI; 1999-494091/41.	
XX		
DR	N-PSDB; AAX87676.	
XX		
PT	Recombinant proteins from recombination of HGF and MSP structural	
PT	domains, useful for protecting cells against apoptosis induced by	
PT	chemotherapeutics.	

This is the amino acid sequence of alphabeta-1 factor, a recombinant protein composed of the hepatocyte growth factor (HGF) alpha chain (1.e., the signal sequence, hairpin loop and King's 1-4), the natural cleavage site of HGF, the macrophage stimulating protein (MSP) beta chain and a polyhistidine tag sequence. This structure allows the recombinant protein to interact with both the HGF receptor (Met) and the MSP receptor (Ron) and thereby induce biological responses which are synergistic and selective compared with the natural factor and truncated forms of the proteins. The portions of DNA encoding the various portions of alphabeta-1 are obtained by PCR amplification of HGF or MSP cDNA and recombined to obtain a hybrid DNA (see AAX876176). Expression vectors, prokaryotic or eukaryotic host cells and a process for preparing recombinant proteins from HGF and MSP are claimed. Alphabeta-1 and other HGF-MSP recombinant proteins are used for the prevention or treatment of chemotherapeutic-induced toxicity such as myelotoxicity, hepatotoxicity, nephrotoxicity, mucotoxicity and neurotoxicity (claimed). They protect cells from death (apoptosis) induced by chemotherapy (e.g., tumors. In particular, they can be used for expansion of marrow precursors, to increase proliferation of the haematopoietic precursors or to stimulate their entry in the cycle. By modification of the proteolytic site, hybrid factors can be obtained which are activated by proteases of the endoplasmic reticulum (such as furines) during their syntheses. When the proteolytic site is removed, permanently immature forms of the factors can be obtained, having a potential partial agonistic or antagonistic activity. Different functional domains can be combined so as to modulate the biological effects

SQ Sequence 729 AA;

Query Match	99.9%	Score 2576;	DB 2;	Length 729;
Best Local Similarity	99.8%	Pred. No. 1.4e-163;		
Matches 441; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

QY ERKRPRNTTTFEPFKSKATLLIKTIDPALIKTKKVVNTADOCANCTRRKGGIPFTCCAKAVPRK 60
Db QRRKNTTTFEPFKSKATLLIKTIDPALIKTKKVVNTADOCANCTRRKGGIPFTCCAKAVPRK 91
QY ARKQCLMFPFNSMSGVYKKEFGHEFDLYENKDIYIRNCIIIGKGRSYGYTISITSGIKCOP 120
Db ARKQCLMFPFNSMSGVYKKEFGHEFDLYENKDIYIRNCIIIGKGRSYGYTISITSGIKCOP 151
QY WSSMTPEHSHYRKGDLQENYCNPRGEEGCMCTFSNPEVRYEVCDIPOCSEVECMTCNG 180

Dd	152	MSMSIIPHEHSYRKRDQIQENYCNPNPGEBSGPMCFISNBEVAYEVODIQCESEVBCMTG	211
Qy	181	ESIRGLMDHTESGKIQRMWDHOTPRHAKFLPERYPDKGPDNNYCRNPDGQRPWCYTLDP	240
Dd	212	ESYRGGLMDHTESGKIQRMWDHOTPRHAKFLPERYPDKGPDNNYCRNPDGQRPWCYTLDP	273
Qy	241	HTREWCALITTCADNTMNDTVPLETTETCIQOQSGYGYNTIINAGIPCCRMBSQYPHE	300
Dd	272	HTREWCALITTCADNTMNDTVPLETTETCIQOQSGYGYNTIINAGIPCCRMBSQYPHE	333
Qy	301	HDMTPENFCKDIRENYCENPDGSESPMCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNG	366
Dd	332	HDMTPENFCKDIRENYCENPDGSESPMCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNG	399
Qy	361	KNYNGNLISQTRSLGTSMDKXMDLHRIIWEPPASKLNNYCRNPDDDAHGPMCTGN	420
Dd	392	KNYNGNLISQTRSLGTSMDKXMDLHRIIWEPPASKLNNYCRNPDDDAHGPMCTGN	453
Qy	421	PLIPMDYCPISRCGDDTPTIIV	442
Dd	452	PLIPMDYCPISRCGDDTPTIIV	473

RESULT 11
ID AAY06622 standard; protein; 729 AA.
XX
XX
AC AAY06622;
XX
DT 26-OCT-1999 (first entry)
XX
DE HGF-MSP hybrid protein alphabeta-RTK factor.
XX
XX Alphabeta-RTK; human; hepatocyte growth factor; HGF; MSP;
KW macrophage stimulating protein; apoptosis; chemotherapy; toxicity
KW therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key
FT Domain
FT 1..483
FT /note= "HGF alpha chain"
FT Peptide
FT 1..31
FT /note= "HSP alpha signal peptide"
FT Protein
FT 32..729
FT /note= "mature protein"
FT Cleavage-site
FT 484..487
FT /note= "furin protease cleavage site"
FT Cleavage-site
FT 488..491
FT /note= "HGF cleavage site"
FT Domain
FT 492..717
FT /note= "MSP beta chain"
FT Peptide
FT 718..729
FT /note= "polystyridine tag"
XX
FN W09338968-A1.
XX
PD 05-AUG-1999.
XX
XX 28-JAN-1999; 99WO-EP000502.
PF
XX 30-JAN-1998; 98IT-MI000180.
PR
XX
PA (DOMP-) DOMPE SPA.
XX
PI Medico E, Michieli P, Collesi C, Caselli G, Comoglio P;
DR WPI: 1999-494091/41.
DR N-PSDB; AAX87677.
XX
XX
PT Recombinant proteins from recombination of HGF and MSP structural
PT domains, useful for protecting cells against apoptosis induced by

PT chemotherapeutics.
XX
XX Example 2; Page 58-60; 63pp; English.
XX
CC This is the amino acid sequence of alphabeta-RTK factor, a recombinant
CC protein composed of the hepatocyte growth factor (HGF) alpha chain (i.e.
CC the signal sequence, hairpin loop and kringles 1-4), a furine protease
CC cleavage site, the natural cleavage site of HGF, the macrophage
CC stimulating protein (MSP) beta chain and a polyhistidine tag sequence.
CC This structure allows the recombinant protein to interact with both the
CC HGF receptor (Met) and the MSP receptor (Ron) and thereby induce
CC biological responses which are synergistic and selective compared with
CC the natural factor and truncated forms of the proteins. The portions of
CC DNA encoding the various portions of alphabeta-RTK are obtained by PCR
CC amplification of HGF or MSP cDNA and recombined to obtain a hybrid DNA
CC (see AAX87677). Expression vectors, prokaryotic or eukaryotic host cells
CC and a process for preparing recombinant proteins from HGF and MSP are
CC claimed. Alphabeta-RTK and other HGF-MSP recombinant proteins are used
CC for the prevention or treatment of chemotherapeutic-induced toxicity such
CC as myelotoxicity, hepatotoxicity, nephrotoxicity, mucotoxicity and
CC neurotoxicity (claimed). They protect cells from death (apoptosis)
CC induced by chemotherapy of, e.g. tumors. In particular, they can be used
CC for expansion of marrow precursors, to increase proliferation of the
CC hematopoietic precursors or to stimulate their entry in the cycle. By
CC modification of the proteolytic site, hybrid factors can be obtained
CC which are activated by proteases of the endoplasmic reticulum (such as
CC furines) during their synthesis. When the proteolytic site is removed,
CC permanently immature forms of the factors can be obtained, having a
CC potential partial agonistic or antagonistic activity. Different
CC functional domains can be combined so as to modulate the biological
CC effects
XX
XX Sequence 729 AA;
SQ

Query Match 99.9%; Score 2576; DB 2; Length 729;

Best Local Similarity 99.8%; Pred. No. 1,4e-163; Mismatches 0; Indels 0; Gaps 0;

Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRRNTIHEFKSAKTLIKIDPALKITKKVNTADOCANRCTRNNKLPPTCAFAVFDK 60
DB :
32 QKRRNTIHEFKSAKTLIKIDPALKITKKVNTADOCANRCTRNNKLPPTCAFAVFDK 91
QY 61 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 120
DB 92 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 151
QY 121 WSSMI.PHEHSYRGKLOENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVECMTCNG 180
DB 152 WSSMI.PHEHSYRGKLOENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVECMTCNG 211
QY 181 ESYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDNVCNRPDQPRPWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDNVCNRPDQPRPWCYTLDP 271
QY 241 HTRWEYCAIKTCADTNDTVPLETTETCIGQGEGYRGTVNTIWNIGIPQGRWDSQYPHE 300
DB 272 HTRWEYCAIKTCADTNDTVPLETTETCIGQGEGYRGTVNTIWNIGIPQGRWDSQYPHE 331
QY 301 HMTPEPNFKCQDLRENYCNRNPDGSGSPWCFPTDNPTRVGCSCQIPRCMDSHODCYRNG 360
DB 332 HMTPEPNFKCQDLRENYCNRNPDGSGSPWCFPTDNPTRVGCSCQIPRCMDSHODCYRNG 391
QY 361 KNYMGSLQTRSGLTCSMDKXNMEIDLRHIFWEPDASKINENVCNRPDDAIGPWCYTGN 420
DB 392 KNYMGSLQTRSGLTCSMDKXNMEIDLRHIFWEPDASKINENVCNRPDDAIGPWCYTGN 451
QY 421 PLIPMDYCPISRCEGDTPTIV 442
DB 452 PLIPMDYCPISRCEGDTPTIV 473

RESULT 12
ADD69672

ID ADD69672 standard; protein, 1109 AA.
XX
XX
AC ADD69672;
XX
DT 15-JAN-2004 (first entry)
XX
DE Chimeric HGF/HBsAg L protein - SEQ ID 20.
XX
KW hollow nanoparticle; HBV surface antigen protein; HBsAg L; hepatotropic;
XX liver disease; gene therapy; chimeric; HGF.
OS Unidentified.
OS Hepatitis B virus.
XX
PN WO2003082345-A1.
XX
PD 09-OCT-2003.
XX
PF 05-MAR-2003; 2003WO-JP002602.
XX
PR 29-MAR-2002; 2002JP-00097280.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PI Kuroda S, Tanizawa K, Kondo A, Ueda M, Sano M, Tada H;
PI WPI, 2003-803976/75.
DR N-PEDB; ADD69671.
XX
DR
XX
PT Drugs comprising protein to form hollow nanoparticles with therapeutic
PT substances encapsulated capable of recognizing target cells or tissues
PT specifically and selectively, applicable in disease treatment.
XX
PS Example F; SEQ ID NO 20; 109pp; Japanese.
XX
CC The invention relates to a novel drug comprising a protein forming hollow
CC nanoparticles and capable of recognising specific cells which is used
CC with therapeutic substances for transfer into cells. The protein is
CC particularly hepatitis B virus (HBV) surface antigen protein (HBsAg). The
CC drugs of the invention demonstrate hepatotropic activity and may be
CC useful for disease treatment, particularly in the therapy of liver
CC diseases. The protein of the invention and subsequently the drug, has
CC selectivity and specificity in targeting cells or tissues for delivering
CC drugs fused with the specific nanoparticle-forming proteins, therefore,
CC gene therapy can be achieved without needing surgery and with reduced
CC side-effects. The current sequence is that of the chimeric HGF/HBsAg L
CC protein (SEQ ID 20) of the invention.
XX
SQ Sequence 1109 AA;

Query Match 99.9%; Score 2576; DB 7; Length 1109;

Best Local Similarity 99.8%; Pred. No. 2.2e-163; Mismatches 0; Indels 0; Gaps 0;

Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRRNTIHEFKSAKTLIKIDPALKITKKVNTADOCANRCTRNNKLPPTCAFAVFDK 60
DB :
418 QKRRNTIHEFKSAKTLIKIDPALKITKKVNTADOCANRCTRNNKLPPTCAFAVFDK 477
QY 61 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 120
DB 478 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 537
QY 121 WSSMI.PHEHSYRGKLOENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVECMTCNG 180
DB 538 WSSMI.PHEHSYRGKLOENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVECMTCNG 597
QY 181 ESYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDNVCNRPDQPRPWCYTLDP 240
DB 598 ESYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDNVCNRPDQPRPWCYTLDP 657
QY 241 HTRWEYCAIKTCADTNDTVPLETTETCIGQGEGYRGTVNTIWNIGIPQGRWDSQYPHE 300
DB 658 HTRWEYCAIKTCADTNDTVPLETTETCIGQGEGYRGTVNTIWNIGIPQGRWDSQYPHE 717

QY 301 HDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGDCYRGN 360
 DB 718 HDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGDCYRGN 777
 QY 361 KNYMGULSOTRSGLTCSMDKXNEDLHRHIFWEPPDASKLNENYCRNPDDAHGPPWCYTGN 420
 DB 778 KNYMGULSOTRSGLTCSMDKXNEDLHRHIFWEPPDASKLNENYCRNPDDAHGPPWCYTGN 837
 QY 421 LIPMDYCPISRCGSDTPTIV 442
 DB 838 LIPMDYCPISRCGSDTPTIV 859

RESULT 13
 AAY57174
 ID AAY57174 standard; protein; 442 AA.
 AC AAY57174;
 XX
 DT 11-FEB-2000 (first entry)
 DE N-terminal region of alpha-chain of HGF mutant (residues 32-478).
 XX
 KM Neovascularization; inhibitor; hepatocyte growth factor; HGF;
 KM cMet-HGF receptor; rheumatoid arthritis; diabetic conjunctivitis;
 KM neonatal conjunctivitis; senile macular degeneration; scarring; mutant;
 KM wound healing; contraceptive; vascularization; Crohn's disease; human.
 OS Synthetic.
 OS Homo sapiens.
 XX
 FT Modified-site 1 Location/Qualifiers
 FT Peptide /note="pyroglutamate"
 FT 130..131
 FT /note="deletion of 5 amino acids (see AAY57173)"
 PN WO955361-A1.
 XX
 PD 04-NOV-1999.
 XX
 PF 06-APR-1999; 99MO-JP001834.
 XX
 PR 28-APR-1998; 98JP-00134681.
 XX
 PA (NAKA/) NAKAMURA T.
 PI Nakamura T;
 XX
 DR WPI; 2000-023263/02.
 XX
 PT Neovascularization inhibitory factors and neovascularization inhibitors,
 PT for preventing or treating, e.g. rheumatoid arthritis and diabetic
 PT conjunctivitis.
 XX
 PS Claim 5; Page 69-72; 78pp; Japanese.
 XX
 CC The invention relates to neovascularization inhibitors comprising a
 CC polypeptide (A) with the amino acid sequence of N-terminal region of
 CC hepatocyte growth factor (HGF), or a polypeptide with an amino acid
 CC sequence derived from (A) which has antagonistic activity against HGF via
 CC the cMet-HGF receptor. The polypeptides can be used as neovascularization
 CC inhibitory factors (e.g. for preventing and treating rheumatoid
 CC arthritis, diabetic conjunctivitis, neonatal conjunctivitis, senile
 CC macular degeneration and excess scarring during wound healing), as a
 CC contraceptive, to modulate abnormal proliferation or vascularization
 CC (e.g. psoriasis, Osler-Webber syndrome, myocardiac angiopoiesis,
 CC peripheral angiectasis, haemophilic rheumatism, angiopoiesis of the eye,
 CC angiofibroma, benign tumor and wound granulation), or to treat or prevent
 CC conditions due to the over-stimulation of endothelial cells (e.g. Crohn's
 CC disease, adenomatous sclerosis, gastrointestinal healing, toughening of
 CC skin and excess scarring). The present sequence represents the N-terminal

CC region of alpha-chain of human HGF with a deletion of 5 amino acids
 CC between residues 130 -131 of the present sequence
 XX
 SQ Sequence 442 AA;
 Query Match 99.8%; Score 2574; DB 3; Length 442;
 Best Local Similarity 100.0%; Pred. No. 1.2e-163;
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RKRRTIHEFKSAKTLIKIDPALIKITKKYNTADQCNRCRNGKLPFTCKAFVFDNA 61
 DB 2 RKRRTIHEFKSAKTLIKIDPALIKITKKYNTADQCNRCRNGKLPFTCKAFVFDNA 61
 QY 62 RKQCLWFPNMSNGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITRSIKQCPW 121
 DB 62 RKQCLWFPNMSNGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITRSIKQCPW 121
 QY 122 SSMIPHEHSYRGKDLQENYCRNPRGEGGPPWCFTSNPEVRYEVCIDIPOCSEVECMTCNGE 181
 DB 122 SSMIPHEHSYRGKDLQENYCRNPRGEGGPPWCFTSNPEVRYEVCIDIPOCSEVECMTCNGE 181
 QY 182 SYRGLMDHTESGKICORMDHOTPHRHKFLPERYPDKGPDNRCRNPDDGPPWCYTLDPH 241
 DB 182 SYRGLMDHTESGKICORMDHOTPHRHKFLPERYPDKGPDNRCRNPDDGPPWCYTLDPH 241
 QY 242 TWMEYCAIKTCADNTMNDTVDPLETTECTIOGQEGYRGVTNINWGIPOCRWDSQYPIEH 301
 DB 242 TWMEYCAIKTCADNTMNDTVDPLETTECTIOGQEGYRGVTNINWGIPOCRWDSQYPIEH 301
 QY 302 DMTPEPFCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGDCYRGN 361
 DB 302 DMTPEPFCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGDCYRGN 361
 QY 362 NYMGULSOTRSGLTCSMDKXNEDLHRHIFWEPPDASKLNENYCRNPDDAHGPPWCYTGN 421
 DB 362 NYMGULSOTRSGLTCSMDKXNEDLHRHIFWEPPDASKLNENYCRNPDDAHGPPWCYTGN 421
 QY 422 LIPMDYCPISRCGSDTPTIV 442
 DB 422 LIPMDYCPISRCGSDTPTIV 442

RESULT 14
 AAR21142
 ID AAR21142 standard; protein; 723 AA.
 XX
 AC AAR21142;
 XX
 DT 25-MAR-2003 (revised)
 DT 20-MAY-1992 (first entry)
 XX
 DE Human TGF- β .
 XX
 KW Tumour cell killing factor; hepatocyte growth factor.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT Region 490..505
 FT /note="Beta-chain N-terminal sequence"
 PN WO9201053-A.
 XX
 PD 23-JAN-1992.
 XX
 PF 13-JUL-1990; 90JP-00185852.
 XX
 PR 13-JUL-1990; 90JP-00185852.
 XX
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.
 XX
 PI Shima N, Higashio K, Nagao M, Oogaki F, Takaoka H, Tsuda E;
 XX

DR WPI; 1992-056868/07.
 DR N-PSDB; AAQ21066.
 XX
 PT Prodn. of liver cell growth factor TCF-II - by culture of a transformant
 PT congl. cdna for tcf-ii and deriving from human foetal lung fibroblast
 PT cells.
 XX
 PS Claim 6; Fig 1; 47pp; Japanese.
 XX
 CC A human foetal lung fibroblast library was screened with a synthetic
 CC probe coding for the N-terminal sequence of TCF-II. A cDNA clone was
 CC isolated and sequenced. The coding sequence can be incorporated into a
 CC plasmid and transformed into a microorganism for production of TCF-II.
 CC The protein itself can be used as a liver cell growth factor or a tumour
 CC cell killing factor. See also AAQ21067-9. (Updated on 25-MAR-2003 to
 CC correct PA field.)
 CC
 XX Sequence 723 AA;
 SQ
 Query Match 99.8%; Score 2573; DB 2; Length 723;
 Best Local Similarity 99.5%; Pred. No. 2,3e-163;
 Matches 440; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EKKRNTTHERKKSAAKTTIKIDPAKIKTKKVNNTADOCANCTRNKGLPTCKAFVFDK 60
 DB :|||||
 DB 32 OKRKNNTIHERKKSAAKTTIKIDPAKIKTKKVNNTADOCANCTRNKGLPTCKAFVFDK 91
 QY 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGVSTITKSGIKOP 120
 DB :|||||
 DB 92 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGVSTITKSGIKOP 151
 QY 121 WSSMTIPEHSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPQCEVECMTCNG 180
 DB :|||||
 DB 122 WSSMTIPEHSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPQCEVECMTCNG 211
 QY 181 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGDDNYCRNPDGQPRPWCYTLDP 240
 DB :|||||
 DB 212 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGDDNYCRNPDGQPRPWCYTLDP 271
 QY 241 HRRWEYCAIKTCADNTMDTVPLETTETCIGOGSGYGTNTVNTWNGIPQCRMSQYRHE 300
 DB :|||||
 DB 272 HRRWEYCAIKTCADNTMDTVPLETTETCIGOGSGYGTNTVNTWNGIPQCRMSQYRHE 331
 QY 301 HDMTEPNFKCDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCMSHGQDCYRNG 360
 DB :|||||
 DB 332 HDMTEPNFKCDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCMSHGQDCYRNG 391
 QY 361 KNYMGNLSOTRSGLTCSMDKXMEDLHRHIFWEPDASKINENYCRNPDGDAHGPCYTCGN 420
 DB :|||||
 DB 392 KNYMGNLSOTRSGLTCSMDKXMEDLHRHIFWEPDASKINENYCRNPDGDAHGPCYTCGN 451
 QY 421 FLIPMDYCPISRCEGDTPTIIV 442
 DB :|||||
 DB 452 FLIPMDYCPISRCEGDTPTIIV 473
 RESULT 15
 AAM59923
 ID AAM59923 standard; protein; 723 AA.
 XX
 AC AAM59923;
 XX
 DT 25-MAR-2003 (revised)
 DT 21-DEC-1998 (first entry)
 XX
 DE Human leukocyte-derived hepatocyte growth factor HCL2.
 XX
 KW Hepatocyte-derived growth factor; HGF; human; liver; hepatoma; diagnosis;
 XX therapy.
 OS Homo sapiens.
 XX
 PN EP859009-A2.

XX
 PD 19-AUG-1998.
 XX
 PF 07-JUN-1991; 98EP-00108130.
 XX
 PR 11-JUN-1990; 90JP-00152474.
 PR 07-JUN-1991; 91EP-00109369.
 XX
 PA (NAKA//) NAKAMURA T.
 XX
 PI Nakamura T, Hagiya M, Seki T, Shimonishi M, Shimizu S, Ihara I;
 PI Sakaguchi M, Aseami O;
 DR WPI; 1998-429650/37.
 DR N-PSDB; AAV53627.
 XX
 PT Preparation of recombinant hepatocyte growth factor polypeptide - by
 PT culturing mammalian cells transformed with vector containing human
 PT leukocyte-derived HGF gene.
 PS
 XX Example 1; Fig 3a-c; 30pp; English.
 CC This is the amino acid sequence for human leukocyte-derived hepatocyte
 CC growth factor (HGF) HLC2, deduced from a cDNA clone (see AAV53627)
 CC obtained from a leukocyte cDNA library. Another leukocyte-derived HGF,
 CC HLC3 (see AAM59924), was identified that differs from HLC3 by having an
 CC extra 5 amino acid residues. A claimed method of producing a HGF
 CC recombinant expression vector that has a promoter (preferably the SV40
 CC early promoter), a gene coding for human leukocyte-derived HGF and the
 CC dihydrofolate reductase gene; (b) culturing the transformed cells in the
 CC presence of successively elevated concentrations of methotrexate; and (c)
 CC recovering the HGF polypeptide from the culture supernatant. HGF
 CC polypeptides may be used as hepatocyte cultivation reagents, liver
 CC regeneration promoters, in basic research on liver function, research on
 CC the action of various hormones and drugs on hepatocytes, research on the
 CC carcinogenesis mechanism of hepatoma, clinical diagnostic reagents using
 CC an antibody against the polypeptide and therapeutic drugs for liver
 CC disease. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR
 CC -2003 to correct PR field.)
 CC
 XX Sequence 723 AA;
 SQ
 Query Match 99.5%; Score 2566; DB 2; Length 723;
 Best Local Similarity 99.3%; Pred. No. 6.6e-163;
 Matches 439; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EKKRNTTHERKKSAAKTTIKIDPAKIKTKKVNNTADOCANCTRNKGLPTCKAFVFDK 60
 DB :|||||
 DB 32 OKRKNNTIHERKKSAAKTTIKIDPAKIKTKKVNNTADOCANCTRNKGLPTCKAFVFDK 91
 QY 61 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGVSTITKSGIKOP 120
 DB :|||||
 DB 92 ARKQCLMPFPNMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGVSTITKSGIKOP 151
 QY 121 WSSMTIPEHSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPQCEVECMTCNG 180
 DB :|||||
 DB 152 WSSMTIPEHSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPQCEVECMTCNG 211
 QY 181 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGDDNYCRNPDGQPRPWCYTLDP 240
 DB :|||||
 DB 212 ESYRGLMDHTESGKICQKWDHQTPIRHKFLPERYPDKGDDNYCRNPDGQPRPWCYTLDP 271
 QY 241 HRRWEYCAIKTCADNTMDTVPLETTETCIGOGSGYGTNTVNTWNGIPQCRMSQYRHE 300
 DB :|||||
 DB 272 HRRWEYCAIKTCADNTMDTVPLETTETCIGOGSGYGTNTVNTWNGIPQCRMSQYRHE 331
 QY 301 HDMTEPNFKCDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCMSHGQDCYRNG 360
 DB :|||||
 DB 332 HDMTEPNFKCDLRENYCRNPDGSESFWCTTDPNIRVGYCSQIPNCMSHGQDCYRNG 391
 QY 361 KNYMGNLSOTRSGLTCSMDKXMEDLHRHIFWEPDASKINENYCRNPDGDAHGPCYTCGN 420
 DB :|||||

1115 Page Blank (uspro)

QY 121 WSSMIPHEHSYRGKDLQENYCNPRGEGGPMCFSTSNPEVREYVCDIPQCSVEECMTGNG 180
DB 152 WSSMIPHEHSYRGKDLQENYCNPRGEGGPMCFSTSNPEVREYVCDIPQCSVEECMTGNG 211
QY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPNPGQPPMPCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPNPGQPPMPCYTLDP 271
QY 241 HTRMEYCAIKTCADNTMNDTVPLETTETCIGQGSGYRGVTNTINNGIPCCQRMDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTETCIGQGSGYRGVTNTINNGIPCCQRMDSQYPHE 331
QY 301 HDMTPENFKCDLRENYCNPRGSGSPMCFSTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 360
DB 332 HDMTPENFKCDLRENYCNPRGSGSPMCFSTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 391
QY 361 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCNPNDDAHGPMCYTGN 420
DB 392 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCNPNDDAHGPMCYTGN 451
QY 421 PLIPMDYCPISRCEGDTTPTIV 442
DB 452 PLIPMDYCPISRCEGDTTPTIV 473

RESULT 2
US-08-605-221-4
; Sequence 4, Application US/08605221
; Publication No. US20030060403A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-B (five amino acids deletion type)
US-08-605-221-4

Query Match 99.9%; Score 2576; DB 8; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.5e-209;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRRNTIHEFKSAKTLIKIDPALKITKKVNTADQCANRCTRNKGLPTCKAFVFDK 60
DB 32 QKRRNTIHEFKSAKTLIKIDPALKITKKVNTADQCANRCTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 120
DB 92 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 151
QY 121 WSSMIPHEHSYRGKDLQENYCNPRGEGGPMCFSTSNPEVREYVCDIPQCSVEECMTGNG 180
DB 152 WSSMIPHEHSYRGKDLQENYCNPRGEGGPMCFSTSNPEVREYVCDIPQCSVEECMTGNG 211
QY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPNPGQPPMPCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPNPGQPPMPCYTLDP 271
QY 241 HTRMEYCAIKTCADNTMNDTVPLETTETCIGQGSGYRGVTNTINNGIPCCQRMDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTETCIGQGSGYRGVTNTINNGIPCCQRMDSQYPHE 331
QY 301 HDMTPENFKCDLRENYCNPRGSGSPMCFSTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 360
DB 332 HDMTPENFKCDLRENYCNPRGSGSPMCFSTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 391
QY 361 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCNPNDDAHGPMCYTGN 420

DB 392 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCNPNDDAHGPMCYTGN 451
QY 421 PLIPMDYCPISRCEGDTTPTIV 442
DB 452 PLIPMDYCPISRCEGDTTPTIV 473

RESULT 3
US-10-509-247A-20
; Sequence 20, Application US/10509247A
; Publication No. US20050181064A1
; GENERAL INFORMATION:
; APPLICANT: Japan Science and Technology Corporation
; APPLICANT: Kuroda, Shunichi
; APPLICANT: Tanizawa, Katsuyuki
; APPLICANT: Kondo, Akihiko
; APPLICANT: Ueda, Masakazu
; APPLICANT: Sano, Masaharu
; APPLICANT: Tada, Hiroko
; TITLE OF INVENTION: DRUG CONTAINING HOLLOW PROTEIN NANOPARTICLES OF PARTICLE-FORMING
; FILE REFERENCE: P023P01/US
; CURRENT APPLICATION NUMBER: US/10/509,247A
; CURRENT FILING DATE: 2004-03-28
; PRIOR APPLICATION NUMBER: PCT/JP03/02602
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: JP2002-97280
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 1109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-509-247A-20

Query Match 99.9%; Score 2576; DB 18; Length 1109;
Best Local Similarity 99.8%; Pred. No. 2.6e-209;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKKRRNTIHEFKSAKTLIKIDPALKITKKVNTADQCANRCTRNKGLPTCKAFVFDK 60
DB 418 QKRRNTIHEFKSAKTLIKIDPALKITKKVNTADQCANRCTRNKGLPTCKAFVFDK 477
QY 61 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 120
DB 478 ARKQCLMPFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGVSTIKSGIKOP 537
QY 121 WSSMIPHEHSYRGKDLQENYCNPRGEGGPMCFSTSNPEVREYVCDIPQCSVEECMTGNG 180
DB 538 WSSMIPHEHSYRGKDLQENYCNPRGEGGPMCFSTSNPEVREYVCDIPQCSVEECMTGNG 597
QY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPNPGQPPMPCYTLDP 240
DB 598 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDNVCNPNPGQPPMPCYTLDP 657
QY 241 HTRMEYCAIKTCADNTMNDTVPLETTETCIGQGSGYRGVTNTINNGIPCCQRMDSQYPHE 300
DB 658 HTRMEYCAIKTCADNTMNDTVPLETTETCIGQGSGYRGVTNTINNGIPCCQRMDSQYPHE 717
QY 301 HDMTPENFKCDLRENYCNPRGSGSPMCFSTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 360
DB 718 HDMTPENFKCDLRENYCNPRGSGSPMCFSTDPNIRVGYCSQIPNCDMSHGQDCYRGNG 777
QY 361 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCNPNDDAHGPMCYTGN 420
DB 778 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNENYCNPNDDAHGPMCYTGN 837
QY 421 PLIPMDYCPISRCEGDTTPTIV 442

Db 838 PLIPMDYCPISRCGDDTPTIV 859

RESULT 4

US-09-951-629-1
Sequence 1, Application US/09951629
Publication No. US20020004480A1

GENERAL INFORMATION:

APPLICANT: NAKAMURA, Toshikazu

TITLE OF INVENTION: Anti-Cancer Agent

FILE REFERENCE: 4373-0101P

CURRENT APPLICATION NUMBER: US/09/951,629

CURRENT FILING DATE: 2001-09-14

PRIOR APPLICATION NUMBER: JP 300728

PRIOR FILING DATE: 1995-10-24

PRIOR APPLICATION NUMBER: PCT/JP96/03105

PRIOR FILING DATE: 1996-10-23

PRIOR APPLICATION NUMBER: US 09/471,032

PRIOR FILING DATE: 1999-12-23

NUMBER OF SEQ ID NOS: 1

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 447

TYPE: PRN

ORGANISM: Homo sapiens

US-09-951-629-1

Query Match 99.4%; Score 2563.5; DB 9; Length 447;

Best Local Similarity 98.7%; Pred. No. 9.9e-209;

Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

1 ERKRRNTIHEFKSAKTLIKIDPALKITKKNVTAADOCANRCTRNGKLPFTCKAFVFDK 60

1 QRRRRNTIHEFKSAKTLIKIDPALKITKKNVTAADOCANRCTRNGKLPFTCKAFVFDK 60

61 ARKQCLMFPNNSSSGVKKEFGHEFDLYENKDYINRCIIIGKRSYKGTYSITSGIKCOP 120

61 ARKQCLMFPNNSSSGVKKEFGHEFDLYENKDYINRCIIIGKRSYKGTYSITSGIKCOP 120

121 WSMIPIHEH-----SYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 175

121 WSMIPIHEHFLPSSTRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 180

176 MTCNGESYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 235

181 MTCNGESYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 240

176 MTCNGESYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 235

181 MTCNGESYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 240

236 YTLDPHTREYCAIKTCADNTAMDTPVPLETTECIGOGEGYRGVNTIINGIIPCORWDS 295

241 YTLDPHTREYCAIKTCADNTAMDTPVPLETTECIGOGEGYRGVNTIINGIIPCORWDS 300

296 QVPEHDMTPENFKCDLRENYCRNPDGSESPMCFITDPIIRGYCSOI PNCMSHGDC 355

301 QVPEHDMTPENFKCDLRENYCRNPDGSESPMCFITDPIIRGYCSOI PNCMSHGDC 360

356 YRGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASLTANENYCRNPDDAHGPW 415

361 YRGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASLTANENYCRNPDDAHGPW 420

416 CYTGNPLIPMDYCPISRCGDDTPTIV 442

421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

CURRENT FILING DATE: 2002-06-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 447
TYPE: PRN
ORGANISM: homosapiens
US-10-081-309-2

Query Match 99.4%; Score 2563.5; DB 14; Length 447;

Best Local Similarity 98.7%; Pred. No. 9.9e-209;

Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

1 ERKRRNTIHEFKSAKTLIKIDPALKITKKNVTAADOCANRCTRNGKLPFTCKAFVFDK 60

1 QRRRRNTIHEFKSAKTLIKIDPALKITKKNVTAADOCANRCTRNGKLPFTCKAFVFDK 60

61 ARKQCLMFPNNSSSGVKKEFGHEFDLYENKDYINRCIIIGKRSYKGTYSITSGIKCOP 120

61 ARKQCLMFPNNSSSGVKKEFGHEFDLYENKDYINRCIIIGKRSYKGTYSITSGIKCOP 120

121 WSMIPIHEH-----SYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 175

121 WSMIPIHEHFLPSSTRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 180

176 MTCNGESYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 235

181 MTCNGESYRGKDLQENYCRNPRGEGGPMCFNSPEVRYEVCIDIPOCSEVBC 240

236 YTLDPHTREYCAIKTCADNTAMDTPVPLETTECIGOGEGYRGVNTIINGIIPCORWDS 295

241 YTLDPHTREYCAIKTCADNTAMDTPVPLETTECIGOGEGYRGVNTIINGIIPCORWDS 300

296 QVPEHDMTPENFKCDLRENYCRNPDGSESPMCFITDPIIRGYCSOI PNCMSHGDC 355

301 QVPEHDMTPENFKCDLRENYCRNPDGSESPMCFITDPIIRGYCSOI PNCMSHGDC 360

356 YRGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASLTANENYCRNPDDAHGPW 415

361 YRGNGKNYGNLSQTRSGLTCSWMDKNMEDLHRHIFWEPDASLTANENYCRNPDDAHGPW 420

416 CYTGNPLIPMDYCPISRCGDDTPTIV 442

421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

Db 421 CYTGNPLIPMDYCPISRCGDDTPTIV 447

```
Db 32 QKRRNTTHERKSAKTLIKIDPALKITKKNVTADOCANRCTRNKGLPFTCKAFVFDK 91
Qy 61 ARKOCMLPFPNMSGSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTVSITKSGIKOP 120
Db 92 ARKOCMLPFPNMSGSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTVSITKSGIKOP 151
Qy 121 WSMIIPHEH-----SYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCIDIPCCSEVEC 175
Db 152 WSMIIPHEHFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCIDIPCCSEVEC 211
Qy 176 MTCGSEYRGKDLMTDESGLICQRMWHDTPHHRKFLPERYPDKGFDNDYCRNPDQPRPWC 235
Db 212 MTCGSEYRGKDLMTDESGLICQRMWHDTPHHRKFLPERYPDKGFDNDYCRNPDQPRPWC 271
Qy 236 YTLDPHTWEXCAIKTCADNTMNDTVPLETTETCICOGGEGYRGTVNTIWMGICQWRDS 295
Db 272 YTLDPHTWEXCAIKTCADNTMNDTVPLETTETCICOGGEGYRGTVNTIWMGICQWRDS 331
Qy 296 QYPHEHMTPENFCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDMSHGQDC 355
Db 332 QYPHEHMTPENFCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDMSHGQDC 391
Qy 356 YRGNGKNYMGNLSTQTRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDQDAHGPW 415
Db 392 YRGNGKNYMGNLSTQTRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDQDAHGPW 451
Qy 416 CYTGNPLIPMDYCPISRCBGDTTPTIV 442
Db 452 CYTGNPLIPMDYCPISRCBGDTTPTIV 478
```

```
RESULT 7
US-10-872-198-120
; Sequence 120, Application US/10872198
; Publication No. US20050002897A1
; GENERAL INFORMATION:
; APPLICANT: Ulrich HAUPTS
; APPLICANT: Andre KOLTERMANN
; APPLICANT: Andreas SCHEIDIG
; APPLICANT: Christian VOETSMER
; APPLICANT: Ulrich Ketting
; TITLE OF INVENTION: NEW BIOLOGICAL ENTITIES AND USE THEREOF
; FILE REFERENCE: 04156.000204
; CURRENT APPLICATION NUMBER: US/10/872,198
; PRIOR FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 60/543,518
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/524,960
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: EP 04003058
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: EP 03025871
; PRIOR FILING DATE: 2003-11-11
; PRIOR APPLICATION NUMBER: EP 03025851
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: EP 03013819
; PRIOR FILING DATE: 2003-06-18
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-872-198-120
```

```
Query Match 99.4%; Score 2563.5; DB 17; Length 726;
Best Local Similarity 98.7%; Pred. No. 1.8e-208;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
```

```
Qy 1 EKKRRNTTHERKSAKTLIKIDPALKITKKNVTADOCANRCTRNKGLPFTCKAFVFDK 60
Db 30 QKRRNTTHERKSAKTLIKIDPALKITKKNVTADOCANRCTRNKGLPFTCKAFVFDK 89
```

```
Qy 61 ARKOCMLPFPNMSGSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTVSITKSGIKOP 120
Db 90 ARKOCMLPFPNMSGSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTVSITKSGIKOP 149
Qy 121 WSMIIPHEH-----SYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCIDIPCCSEVEC 175
Db 150 WSMIIPHEHFLPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCIDIPCCSEVEC 209
Qy 176 MTCGSEYRGKDLMTDESGLICQRMWHDTPHHRKFLPERYPDKGFDNDYCRNPDQPRPWC 235
Db 210 MTCGSEYRGKDLMTDESGLICQRMWHDTPHHRKFLPERYPDKGFDNDYCRNPDQPRPWC 269
Qy 236 YTLDPHTWEXCAIKTCADNTMNDTVPLETTETCICOGGEGYRGTVNTIWMGICQWRDS 295
Db 270 YTLDPHTWEXCAIKTCADNTMNDTVPLETTETCICOGGEGYRGTVNTIWMGICQWRDS 329
Qy 296 QYPHEHMTPENFCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDMSHGQDC 355
Db 330 QYPHEHMTPENFCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDMSHGQDC 389
Qy 356 YRGNGKNYMGNLSTQTRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDQDAHGPW 415
Db 390 YRGNGKNYMGNLSTQTRSGLTCSMDKXMEDLHRHIFWEPDASKLNENYCRNPDQDAHGPW 449
Qy 416 CYTGNPLIPMDYCPISRCBGDTTPTIV 442
Db 450 CYTGNPLIPMDYCPISRCBGDTTPTIV 476
```

```
RESULT 8
US-11-021-951-120
; Sequence 120, Application US/11021951
; Publication No. US20050175581A1
; GENERAL INFORMATION:
; APPLICANT: HAUPTS, Ulrich
; APPLICANT: KOLTERMANN, Andre
; APPLICANT: SCHEIDIG, Andreas
; APPLICANT: VOETSMER, Christian
; APPLICANT: Ketting, Ulrich
; APPLICANT: COCO, Wayne Michael
; TITLE OF INVENTION: New Biological Entities And The Pharmaceutical
; TITLE OF INVENTION: And Diagnostic Use Thereof
; FILE REFERENCE: 04156.000205
; CURRENT APPLICATION NUMBER: US/11/021,951
; PRIOR FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: 10/872,198
; PRIOR FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 60/543,518
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/524,960
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: EP 04003058
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: EP 03025871
; PRIOR FILING DATE: 2003-11-11
; PRIOR APPLICATION NUMBER: EP 03025851
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: EP 03013819
; PRIOR FILING DATE: 2003-06-18
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-021-951-120
```

```
Query Match 99.4%; Score 2563.5; DB 20; Length 726;
Best Local Similarity 98.7%; Pred. No. 1.8e-208;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
```

```
Qy 1 EKKRRNTTHERKSAKTLIKIDPALKITKKNVTADOCANRCTRNKGLPFTCKAFVFDK 60
Db 30 QKRRNTTHERKSAKTLIKIDPALKITKKNVTADOCANRCTRNKGLPFTCKAFVFDK 89
```

Db 30 QRRRRTTIEFKSAKTTLIKIDPALKITKKNVTADQCANRCTRNGGLPFTCKAFVFDK 89
QY 61 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 120
Db 90 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 149
QY 121 WSSMIPIHEH-----SYRGKDLOENYCRNPRGEGGFWCFTSNPEVRYEVCDIPOCSEVBC 175
Db 150 WSSMIPIHEHFLPSSYRGKDLOENYCRNPRGEGGFWCFTSNPEVRYEVCDIPOCSEVBC 209
QY 176 MTCNGESYGLMDHTSNGIKICQRMWDQTPRHRKFLPERYDYGFDNNYCRNPDGQRPWC 235
Db 210 MTCNGESYGLMDHTSNGIKICQRMWDQTPRHRKFLPERYDYGFDNNYCRNPDGQRPWC 269
QY 236 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIWNGLPCORWDS 295
Db 270 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIWNGLPCORWDS 329
QY 296 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQI.PNCDSHGQDC 355
Db 330 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQI.PNCDSHGQDC 389
QY 356 YRGNGKNYMGNTLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLBNENYCRNPDDAHGPW 415
Db 390 YRGNGKNYMGNTLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLBNENYCRNPDDAHGPW 449
QY 416 CYTGNPLIPWDYCPISRCBGDTTPTIV 442
Db 450 CYTGNPLIPWDYCPISRCBGDTTPTIV 476

RESULT 9

US-08-605-221-2
; Sequence 2, Application US/08605221
; Publication No. US20030060403A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match 99.4%; Score 2563.5; DB 8; Length 728;
Best Local Similarity 98.7%; Pred. No. 1.8e-208;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
QY 1 ERKRRNTTIEFKSAKTTLIKIDPALKITKKNVTADQCANRCTRNGGLPFTCKAFVFDK 60
Db 32 QRRRRTTIEFKSAKTTLIKIDPALKITKKNVTADQCANRCTRNGGLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 120
Db 92 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 151
QY 121 WSSMIPIHEH-----SYRGKDLOENYCRNPRGEGGFWCFTSNPEVRYEVCDIPOCSEVBC 175
Db 152 WSSMIPIHEHFLPSSYRGKDLOENYCRNPRGEGGFWCFTSNPEVRYEVCDIPOCSEVBC 211
QY 176 MTCNGESYGLMDHTSNGIKICQRMWDQTPRHRKFLPERYDYGFDNNYCRNPDGQRPWC 235
Db 212 MTCNGESYGLMDHTSNGIKICQRMWDQTPRHRKFLPERYDYGFDNNYCRNPDGQRPWC 271
QY 236 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIWNGLPCORWDS 295
Db 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIWNGLPCORWDS 331

QY 296 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQI.PNCDSHGQDC 355
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQI.PNCDSHGQDC 391
QY 356 YRGNGKNYMGNTLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLBNENYCRNPDDAHGPW 415
Db 392 YRGNGKNYMGNTLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLBNENYCRNPDDAHGPW 451
QY 416 CYTGNPLIPWDYCPISRCBGDTTPTIV 442
Db 452 CYTGNPLIPWDYCPISRCBGDTTPTIV 478

RESULT 10

US-10-813-805-10
; Sequence 10, Application US/10813805
; Publication No. US20040191902A1
; GENERAL INFORMATION:
; APPLICANT: PEIZER INC.
; APPLICANT: Hamor, John E.
; APPLICANT: Roach, Marsha L.
; TITLE OF INVENTION: GROWTH AND DIFFERENTIATION OF STEM CELLS
; FILE REFERENCE: PC25028A
; CURRENT APPLICATION NUMBER: US/10/813,805
; CURRENT FILING DATE: 2004-03-30
; PRIOR APPLICATION NUMBER: US 60/459,449
; PRIOR FILING DATE: 2003-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 10
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-813-805-10

Query Match 99.4%; Score 2563.5; DB 16; Length 728;
Best Local Similarity 98.7%; Pred. No. 1.8e-208;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
QY 1 ERKRRNTTIEFKSAKTTLIKIDPALKITKKNVTADQCANRCTRNGGLPFTCKAFVFDK 60
Db 32 QRRRRTTIEFKSAKTTLIKIDPALKITKKNVTADQCANRCTRNGGLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 120
Db 92 ARKQCLMPFPNSMSGVKKFEGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 151
QY 121 WSSMIPIHEH-----SYRGKDLOENYCRNPRGEGGFWCFTSNPEVRYEVCDIPOCSEVBC 175
Db 152 WSSMIPIHEHFLPSSYRGKDLOENYCRNPRGEGGFWCFTSNPEVRYEVCDIPOCSEVBC 211
QY 176 MTCNGESYGLMDHTSNGIKICQRMWDQTPRHRKFLPERYDYGFDNNYCRNPDGQRPWC 235
Db 212 MTCNGESYGLMDHTSNGIKICQRMWDQTPRHRKFLPERYDYGFDNNYCRNPDGQRPWC 271
QY 236 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIWNGLPCORWDS 295
Db 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIWNGLPCORWDS 331
QY 296 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQI.PNCDSHGQDC 355
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQI.PNCDSHGQDC 391
QY 356 YRGNGKNYMGNTLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLBNENYCRNPDDAHGPW 415
Db 392 YRGNGKNYMGNTLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLBNENYCRNPDDAHGPW 451
QY 416 CYTGNPLIPWDYCPISRCBGDTTPTIV 442
Db 452 CYTGNPLIPWDYCPISRCBGDTTPTIV 478

RESULT 11
US-10-475-616-2
; Sequence 2, Application US/10475616
; Publication No. US20040236073A1
; GENERAL INFORMATION:
; APPLICANT: Gherard, Ermanno
; APPLICANT: Liebha, Daniel
; APPLICANT: Blundell, Thomas L
; APPLICANT: Chigade, Dmitry Y
; TITLE OF INVENTION: The NK1 Fragment of Hepatocyte Growth Factor/Scatter
; TITLE OF INVENTION: The NK1 Fragment of Hepatocyte Growth Factor/Scatter
; FILE REFERENCE: 2502489-991180
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: PCT/GB02/01941
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: GB 0110430.6
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-475-616-2

Query Match 99.4%; Score 2563.5; DB 16; Length 728;
Best Local Similarity 98.7%; Pred. No. 1.8e-208;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVBDK 60
DB :
QY 32 QKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVBDK 91
DB :
QY 61 ARKQCLMPFPMSSSGVKKFEGHFDLYENKDIYRNCIIIGKRSYKGVSTIKSGIKOP 120
DB 92 ARKQCLMPFPMSSSGVKKFEGHFDLYENKDIYRNCIIIGKRSYKGVSTIKSGIKOP 151
QY 121 WSSMI.PHEH-----SYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCIDIPQCEVEEC 175
DB 152 WSSMI.PHEHSLPSPSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCIDIPQCEVEEC 211
QY 176 MTCNGESYRGLMDHTESGKICQRMWHQTPHKKFLPERYPDXGFDNNYCRNPDQOPRWC 235
DB 212 MTCNGESYRGLMDHTESGKICQRMWHQTPHKKFLPERYPDXGFDNNYCRNPDQOPRWC 271
QY 236 YTLDPHTHWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIANGIPQORWDS 295
DB 272 YTLDPHTHWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIANGIPQORWDS 331
QY 296 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHGQDC 355
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 356 YRGNKNYMGSLQSTRSGLTCSMDKXMDLHRHIFWEPDASKLNENYCRNPDHAGPW 415
DB 392 YRGNKNYMGSLQSTRSGLTCSMDKXMDLHRHIFWEPDASKLNENYCRNPDHAGPW 451
QY 416 CYTGNPLIPWDYCPISRCEGDTPTIV 442
DB 452 CYTGNPLIPWDYCPISRCEGDTPTIV 478

RESULT 12
US-10-498-332-7
; Sequence 7, Application US/10498332
; Publication No. US20050113284A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Motoao
; APPLICANT: Higuchi, Toshio
; APPLICANT: Yamasaki, Yoshiaki
; APPLICANT: Orita, Takuya
; TITLE OF INVENTION: PHARMACEUTICAL AGENTS AND METHODS FOR

; TITLE OF INVENTION: TISSUE AND VASCULAR REGENERATION
; FILE REFERENCE: 14539-016US1
; CURRENT APPLICATION NUMBER: US/10/498.332
; CURRENT FILING DATE: 2004-06-11
; PRIOR APPLICATION NUMBER: PCT/JP02/13014
; PRIOR FILING DATE: 2002-12-12
; PRIOR APPLICATION NUMBER: JP 2002-352924
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: JP 2001-380158
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-498-332-7

Query Match 99.4%; Score 2563.5; DB 17; Length 728;
Best Local Similarity 98.7%; Pred. No. 1.8e-208;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVBDK 60
DB :
QY 32 QKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPTCKAFVBDK 91
DB :
QY 61 ARKQCLMPFPMSSSGVKKFEGHFDLYENKDIYRNCIIIGKRSYKGVSTIKSGIKOP 120
DB 92 ARKQCLMPFPMSSSGVKKFEGHFDLYENKDIYRNCIIIGKRSYKGVSTIKSGIKOP 151
QY 121 WSSMI.PHEH-----SYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCIDIPQCEVEEC 175
DB 152 WSSMI.PHEHSLPSPSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCIDIPQCEVEEC 211
QY 176 MTCNGESYRGLMDHTESGKICQRMWHQTPHKKFLPERYPDXGFDNNYCRNPDQOPRWC 235
DB 212 MTCNGESYRGLMDHTESGKICQRMWHQTPHKKFLPERYPDXGFDNNYCRNPDQOPRWC 271
QY 236 YTLDPHTHWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIANGIPQORWDS 295
DB 272 YTLDPHTHWEYCAIKTCADNTMNDTVPLETTTCIOGQSGYRGVTNTIANGIPQORWDS 331
QY 296 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHGQDC 355
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHGQDC 391
QY 356 YRGNKNYMGSLQSTRSGLTCSMDKXMDLHRHIFWEPDASKLNENYCRNPDHAGPW 415
DB 392 YRGNKNYMGSLQSTRSGLTCSMDKXMDLHRHIFWEPDASKLNENYCRNPDHAGPW 451
QY 416 CYTGNPLIPWDYCPISRCEGDTPTIV 442
DB 452 CYTGNPLIPWDYCPISRCEGDTPTIV 478

RESULT 13
US-10-133-912-18
; Sequence 18, Application US/10133912
; Publication No. US20020165358A1
; GENERAL INFORMATION:
; APPLICANT: Kinoshita, Masahiko
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Goto, Masaki
; APPLICANT: Murakami, Akihiko
; APPLICANT: Ueda, Maatsugu
; APPLICANT: Higashio, Kanji
; APPLICANT: Yamashita, Yasushi
; TITLE OF INVENTION: TCF MUTANT
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK

STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/5 inch diskette
COMPUTER: PC'S LIMITED SYSTEM 200
OPERATING SYSTEM: WORD FOR WINDOWS
SOFTWARE: MICROSOFT WINDOWS 98
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/133,912
FILING DATE: 25-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/700,519
FILING DATE: 26-Aug-1996
ATTORNEY/AGENT INFORMATION:
NAME: Wayne, Milton J.
REGISTRATION NUMBER: 17,906
REFERENCE/DOCKET NUMBER: U-Wp-5240PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-683-8150
TELEFAX: 212-532-4285
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 723
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-133-912-18

Query Match 99.0%; Score 2554; DB 13; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.1e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 NTHERKSKAKTLIKIDPALKIKTKKVNADOCANRCTRNKGLPTCKAFVFDKARKQC 65
DB 37 NTHERKSKAKTLIKIDPALKIKTKKVNADOCANRCTRNKGLPTCKAFVFDKARKQC 96
QY 66 LMFPPMSMSGVKEFGHEFDLYENKDYINNCIIIGKRSYKGTVSITKSGIKCOPWSSMT 125
DB 97 LMFPPMSMSGVKEFGHEFDLYENKDYINNCIIIGKRSYKGTVSITKSGIKCOPWSSMT 156
QY 126 PHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDDPQCEVEVEMTCNGESYRG 185
DB 157 PHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDDPQCEVEVEMTCNGESYRG 216
QY 186 LMDHTSSGKICQWMDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLLDPHTRWE 245
DB 217 LMDHTSSGKICQWMDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLLDPHTRWE 276
QY 246 YCAIKTCADNTMDTVPLETTTCIOGQGGYGTVNTIWNGLPCORWDSQYFHEHDMTP 305
DB 277 YCAIKTCADNTMDTVPLETTTCIOGQGGYGTVNTIWNGLPCORWDSQYFHEHDMTP 336
QY 306 ENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDCYRNGNGXNYG 365
DB 337 ENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDCYRNGNGXNYG 396
QY 366 NLSQTSGLTCSMWDKXMEDLHRHITWEPDASKLNNYCRNPDDAHGPWCYTGNPLIWM 425
DB 397 NLSQTSGLTCSMWDKXMEDLHRHITWEPDASKLNNYCRNPDDAHGPWCYTGNPLIWM 456
QY 426 DYCPIRSRGGDTPTTIV 442
DB 457 DYCPIRSRGGDTPTTIV 473

RESULT 14
US-10-133-912-19
Sequence 19: Application US/10133912
Publication NO: US20020165358A1
GENERAL INFORMATION:

APPLICANT: Kinoshaki, Masahiko
Yamaguchi, Kyoji
Goto, Masaki
Murakami, Akihiko
Ueda, Masatsugu
Higashio, Kanji
Yamashita, Yasushi
TITLE OF INVENTION: TCF MUTANT
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burgess, Ryan and Wayne
STREET: 370 Lexington Avenue, Suite 2105
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/5 inch diskette
COMPUTER: PC'S LIMITED SYSTEM 200
OPERATING SYSTEM: WORD FOR WINDOWS
SOFTWARE: MICROSOFT WINDOWS 98
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/133,912
FILING DATE: 25-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/700,519
FILING DATE: 26-Aug-1996
ATTORNEY/AGENT INFORMATION:
NAME: Wayne, Milton J.
REGISTRATION NUMBER: 17,906
REFERENCE/DOCKET NUMBER: U-Wp-5240PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-683-8150
TELEFAX: 212-532-4285
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 723
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-133-912-19

Query Match 99.0%; Score 2552; DB 13; Length 723;
Best Local Similarity 98.9%; Pred. No. 1.7e-207;
Matches 437; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 ERKRNTIHEFKSKAKTLIKIDPALKIKTKKVNADOCANRCTRNKGLPTCKAFVFDK 60
DB 32 ERKRNTIHEFKSKAKTLIKIDPALAATAVNTADOCANRCTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLMPFPNMSGVKEFGHEFDLYENKDYINNCIIIGKRSYKGTVSITKSGIKCOP 120
DB 92 ARKQCLMPFPNMSGVKEFGHEFDLYENKDYINNCIIIGKRSYKGTVSITKSGIKCOP 151
QY 121 WSSMT PHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDDPQCEVEVEMTCNG 180
DB 152 WSSMT PHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDDPQCEVEVEMTCNG 211
QY 181 ESYRGLMDHTSSGKICQWMDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLLDP 240
DB 212 ESYRGLMDHTSSGKICQWMDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLLDP 271
QY 241 HTRWEYCAIKTCADNTMDTVPLETTTCIOGQGGYGTVNTIWNGLPCORWDSQYFHE 300
DB 272 HTRWEYCAIKTCADNTMDTVPLETTTCIOGQGGYGTVNTIWNGLPCORWDSQYFHE 331
QY 301 HDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDCYRNGNG 360
DB 332 HDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHGQDCYRNGNG 391
QY 361 KNYMGNLSQTSGLTCSMWDKXMEDLHRHITWEPDASKLNNYCRNPDDAHGPWCYTGN 420

Db 392 KVMNLSQTRSGLTCSMWKMDKMDLHRHIFWEPDASKLINENYCRNPPDDAHGPMCYTGN 451
QY 421 PLIPWDYCPISRCEGDTTPTIV 442
Db 452 PLIPWDYCPISRCEGDTTPTIV 473

RESULT 15

US-10-149-103A-1
; Sequence 1, Application US/10149103A
; Publication No. US20040121945A1
; GENERAL INFORMATION:
; APPLICANT: Liang, Hong
; APPLICANT: Sim, Kim Lee
; APPLICANT: Day, Clara
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0592 43170-274540
; CURRENT APPLICATION NUMBER: US/10/149,103A
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/US00/34039
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-149-103A-1

Query Match 98.0%; Score 2527.5; DB 16; Length 727;
Best Local Similarity 97.1%; Pred. No. 2e-205;
Matches 433; Conservative 6; Mismatches 2; Indels 5; Gaps 1;

QY 2 RRRRTTHFEKKSATKTLIKIDPALKIKTKKYNADQCANRCTRNKGLPFTCKAFVFDKA 61
Db 33 KRRRTTHFEKKSATKTLIKIDPALKIKTKKYNADQCANRCTRNKGLPFTCKAFVFDKA 92
QY 62 RQCLMFPPNSMSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVAITKSGIKQCPW 121
Db 93 RQCLMFPPNSMSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVAITKSGIKQCPW 152
QY 122 SSMIPHEH-----SYRGKDLQENYCRNPRGEGGPMCFTSNPEVRYEVCDIPOCSEVECM 176
Db 153 SSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGPMCFTSNPEVRYEVCDIPOCSEVECM 212
QY 177 TNGBSYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDNYCRNPDGQPRPMCY 236
Db 213 TNGBSYRGLMDHTESGKICQRMWDHOTPHRHKFLPERYPDKGFDNYCRNPDGQPRPMCY 272
QY 237 TLDPHTREYCAIKTCADNTMNDTVPLTTECIQOGGEGYGTVNTINNGIPCORWDSQ 296
Db 273 TLDPHTREYCAIKTCADNTMNDTVPLTTECIQOGGEGYGTVNTINNGIPCORWDSQ 332
QY 297 YPHEHDMTPENFKCXDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCMSHGQDCY 356
Db 333 YPHKIDMTPENFCKXDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCMSHGQDCY 392
QY 357 RGNKKNYMGNLISQTRSGLTCSMWKMDKMDLHRHIFWEPDASKLINENYCRNPPDDAHGPMW 416
Db 393 RGNKKNYMGNLISQTRSGLTCSMWKMDKMDLHRHIFWEPDASKLINENYCRNPPDDAHGPMW 452
QY 417 YTGNPPLIPWDYCPISRCEGDTTPTIV 442
Db 453 YTGNPPLIPWDYCPISRCEGDTTPTIV 478

Search completed: October 3, 2005, 07:42:28
Job time : 239.645 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2005, 05:58:32 ; Search time 1360.31 Seconds
(without alignments)
24.255 Million cell updates/sec

Title: US-09-674-377B-2

Perfect score: 2579
Sequence: 1 EKKRRNTIHEFKSAKTLLI.....IPWDYCPISRCBGDTTPRIV 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2576	99.9	487	4	US-09-601-040A-27
2	2576	99.9	717	4	US-09-601-040A-6
3	2576	99.9	717	4	US-09-601-040A-8
4	2576	99.9	723	1	US-07-838-410-1
5	2576	99.9	723	1	US-08-290-937B-1
6	2576	99.9	723	1	US-08-290-937B-2
7	2576	99.9	723	1	US-08-404-643-1
8	2576	99.9	723	3	US-09-194-326-1
9	2576	99.9	723	4	US-09-600-991-18
10	2576	99.9	723	4	US-08-605-221-4
11	2576	99.9	723	4	US-09-601-040A-10
12	2576	99.9	729	4	US-09-601-040A-2
13	2576	99.9	729	4	US-09-601-040A-4
14	2573	99.8	723	3	US-08-030-810-3
15	2570	99.7	723	3	US-08-290-937B-3
16	2563.5	99.4	697	2	US-08-460-890A-50
17	2563.5	99.4	697	2	US-08-167-641C-50
18	2563.5	99.4	697	3	US-08-460-971A-50
19	2563.5	99.4	697	3	US-08-462-040-50
20	2563.5	99.4	728	1	US-07-815-333A-2
21	2563.5	99.4	728	4	US-08-087-783A-22
22	2563.5	99.4	728	4	US-08-605-221-2
23	2554	99.0	723	3	US-09-194-326-2
24	2554	99.0	723	3	US-08-700-519D-18
25	2552	99.0	723	3	US-09-194-326-3
26	2552	99.0	723	3	US-08-700-519D-19
27	1793	69.5	574	4	US-09-600-991-2

28	1460	56.6	563	4	US-09-600-991-4	Sequence 4, Appli
29	1433.5	55.6	290	4	US-08-484-841A-8	Sequence 8, Appli
30	1171	45.4	711	1	US-08-184-012C-8	Sequence 8, Appli
31	1171	45.4	711	1	US-08-334-177-2	Sequence 2, Appli
32	1171	45.4	711	2	US-08-666-082B-1	Sequence 1, Appli
33	1171	45.4	711	2	US-08-766-982-2	Sequence 2, Appli
34	1171	45.4	711	3	US-09-296-219-2	Sequence 2, Appli
35	1171	45.4	711	4	US-09-600-991-20	Sequence 20, Appli
36	1171	45.4	711	4	US-09-601-040A-12	Sequence 12, Appli
37	1171	45.4	711	4	US-09-949-016-6981	Sequence 6981, Ap
38	1171	45.4	711	5	PCT-US95-13830-2	Sequence 2, Appli
39	1114.5	43.2	716	2	US-08-766-982-1	Sequence 1, Appli
40	1114.5	43.2	716	3	US-09-296-219-1	Sequence 1, Appli
41	1054.5	40.9	790	4	US-08-991-761A-13	Sequence 13, Appli
42	1039.5	40.3	812	1	US-08-248-629A-1	Sequence 1, Appli
43	1039.5	40.3	812	1	US-08-451-932-1	Sequence 1, Appli
44	1039.5	40.3	812	1	US-08-452-260-1	Sequence 1, Appli
45	1039.5	40.3	812	1	US-08-326-785-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-601-040A-27
; Sequence 27, Application US/09601040A
; Patent No. 6730657
; GENERAL INFORMATION:
; APPLICANT: MEDICO, Enzo et al.
; TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
; FILE REFERENCE: 0471-0161P
; CURRENT APPLICATION NUMBER: US/09/601,040A
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-601-040A-27
```

Query Match 99.9%; Score 2576; DB 4; Length 487;

Best Local Similarity 99.8%; Pred. No. 1.1e-213;

Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 EKKRRNTIHEFKSAKTLLIKITDPALKIKTKKVTNADOCANRCTNKGLPFTCKAFVFXK 60
DB 32 QKKRRNTIHEFKSAKTLLIKITDPALKIKTKKVTNADOCANRCTNKGLPFTCKAFVFXK 91
QY 61 ARKQCIWPFPPNMSGSKVKEFGHEFDLYENKDIYINNCIIIGKRSYKGVSIITKSGIKCOP 120
DB 92 ARKQCIWPFPPNMSGSKVKEFGHEFDLYENKDIYINNCIIIGKRSYKGVSIITKSGIKCOP 151
QY 121 WSSMTIPHEHSYKGLQENYCNPRGEGGPMCFPSNPEVYECVDIPCCSEVECMTCNG 180
DB 152 WSSMTIPHEHSYKGLQENYCNPRGEGGPMCFPSNPEVYECVDIPCCSEVECMTCNG 211
QY 181 ESYRGIMDTESGKICQRMHDQTPRHKFLPRRYDCKGDDVYCNRPDQPPMCTYLPD 240
DB 212 ESYRGIMDTESGKICQRMHDQTPRHKFLPRRYDCKGDDVYCNRPDQPPMCTYLPD 271
QY 241 HTRWEYCAIKTCADYTMNDTVPLETTETCIQCGEGYRGYTVTINNGIPQRMDSQYPHE 300
DB 272 HTRWEYCAIKTCADYTMNDTVPLETTETCIQCGEGYRGYTVTINNGIPQRMDSQYPHE 331
QY 301 HMTPEINFCKDLRENYCNPDGSSPMCFITDPINRGYCGQIYINCDMSHQDCYRNG 360
DB 332 HMTPEINFCKDLRENYCNPDGSSPMCFITDPINRGYCGQIYINCDMSHQDCYRNG 391
QY 361 KYYMGVLSOTRSGGLCSMWDKXMDLHRIIYFEPASKLINENYCNRPDDAGPWCYTGN 420
DB 392 KYYMGVLSOTRSGGLCSMWDKXMDLHRIIYFEPASKLINENYCNRPDDAGPWCYTGN 451
```

```

QY      421 PLIPWDYCPISRCEGDTPTIV 442
         |||||
Db      452 PLIPWDYCPISRCEGDTPTIV 473

```

RESULT 2

```

US-09-601-040A-6
? Sequence 6, Application US/09601040A
? Patent No. 6730657
? GENERAL INFORMATION:
? APPLICANT: MEDICO, Enzo et al.
? TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
? FILE REFERENCE: 0471-0161P
? CURRENT APPLICATION NUMBER: US/09/601,040A
? CURRENT FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 28
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 6
? LENGTH: 717
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Recombinant sequence derived from Homo sapiens
? US-09-601-040A-6

```

Query Match	99.9%	Score 2576;	DB 4;	Length 717;
Best Local Similarity	99.8%;	Pred. No.1.8e-213;		
Matches 441;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0

Qy	ERKRNRNIHEFFKSAKTLTKIDPAKI	KTKKVNTADQCANCRNKLPPTCAFAFYDK	60
Db	32	QRRGRNIIHEFFKSAKTLTKIDPAKI	KTKKVNTADQCANCRNKLPPTCAFAFYDK 91
Qy	61	ARKOCLMPFNSSSGYKKEFGHEPDL	YENKDYINNCIIIGKRSYKGVSIITKSGIKOP 120
Db	92	ARKOCLMPFNSSSGYKKEFGHEPDL	YENKDYINNCIIIGKRSYKGVSIITKSGIKOP 151
Qy	121	WSSMIPIHEHSYRGDIOENYCRNPR	BEGBPMCFINSPEVREYVDIQOSEVEOMONG 180
Db	152	WSSMIPIHEHSYRGDIOENYCRNPR	BEGBPMCFINSPEVREYVDIQOSEVEOMONG 211
Qy	181	ESYGLMDHTESGKICORMDQHPHAK	LPERYPDKGFDMDNYCNPDOERPMCYTLDP 240
Db	212	ESYGLMDHTESGKICORMDQHPHAK	LPERYPDKGFDMDNYCNPDOERPMCYTLDP 271
Qy	241	HTREYCAITKCADNMTNDVLETT	ECIOGSEBYRGTVNTINNGIPQCRMSQYPHE 300
Db	272	HTREYCAITKCADNMTNDVLETT	ECIOGSEBYRGTVNTINNGIPQCRMSQYPHE 331
Qy	301	HDMTPENFKCDJRENYCRNPDS	SPMCFETTDPNIRVYCSQIPNCMSHGODCYRNG 360
Db	332	HDMTPENFKCDJRENYCRNPDS	SPMCFETTDPNIRVYCSQIPNCMSHGODCYRNG 391
Qy	361	KNYVGNLSQTRSGJITCSMDKN	MEDLHRHIFWEPDASKLINENYCRNPDDAHGWCYTYGN 420
Db	392	KNYVGNLSQTRSGJITCSMDKN	MEDLHRHIFWEPDASKLINENYCRNPDDAHGWCYTYGN 451
Qy	421	PLIPMDYCPISRCGDTPTTIV	442
Db	452	PLIPMDYCPISRCGDTPTTIV	473

RESULT 3

```

US-09-601-040A-8
Sequence 8, Application US/09601040A
Patent No. 6,730,657
GENERAL INFORMATION:
APPLICANT: MEDICO, Enzo et al.
TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
FILE REFERENCE: 0471-0161P
CURRENT APPLICATION NUMBER: US/09/601,040A
CURRENT FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 28

```

```

: SOFTWARE: Patentin version 3.0
: SEQ ID NO 8
: LENGTH: 717
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Recombinant sequence derived from Homo sapiens
: US-09-601-040A-8

```

Query Match	99.9%	Score 2576;	DB 4;	Length 717;
Best Local Similarity	99.8%	Pred. No. 1.8e-213;		
Matches 441; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	ERRRRTTHEFFKSKAKTLLIKIDPALKIKTKKXNNTADOCANRCITRNGLPEPTCAFYEDK	60
Dp	32	ÖRRRRRTTHEFFKSATLLIKIDPALKIKTKKXNNTADOCANRCITRNGLPEPTCAFYDK	91
Qy	61	ARKQCLMPFPNMSGSGVKKEFGHEFDLYENKDIYRNCIIGKGRSYKGTSLITSKGIKOP	120
Dp	92	ARKQCLMPFPNMSGSGVKKEFGHEFDLYENNDIYRNCIIGKGRSYKGTSLITSKGIKOP	151
Qy	121	WSSMILPHEHSYRGKDLOENYCRNPBGEGGWCFTSNPEVRYEYCDIPQSEVECMTCNG	180
Dp	152	WSSMILPHEHSYRGKDLOENYCRNPBGEGGWCFTSNPEVRYEYCDIPQSEVECMTCNG	211
Qy	181	ESYRGIMDHTESGKICORMDHOTPRRHHFLPERYDYDKDFDNYCRNPDGPRMPCYTLDP	240
Dp	212	ESYRGIMDHTESGKICORMDHOTPRRHHFLPERYDYDKDFDNYCRNPDGPRMPCYTLDP	271
Qy	241	HTRWEXCAIKTCADNTMNDTDVPLETTECICOGQEGYGVNTIWNIGPCORMDSQYPHE	300
Dp	272	HTRWEXCAIKTCADNTMNDTDVPLETTECICOGQEGYGVNTIWNIGPCORMDSQYPHE	331
Qy	301	HDMTPENFCKDLRENYCRNPBGSSPMCFITDPDIRYGVCSOIPNCDMSHGDCYVANG	360
Dp	332	HDMTPENFCKDLRENYCRNPBGSSPMCFITDPDIRYGVCSOIPNCDMSHGDCYVANG	391
Qy	361	KNTMGNLSQTRSGLTCSMWDKXMBELHRIHFWEPDASKLNIENYCRNPDDAHGWCYTCGN	420
Dp	392	KNTMGNLSQTRSGLTCSMWDKXMBELHRIHFWEPDASKLNIENYCRNPDDAHGWCYTCGN	451
Qy	421	PLIPMDYCPISRCBGDTTPTIV 442	
Dp	452	PLIPMDYCPISRCBGDTTPTIV 473	

RESULT 4

```

: Sequence 1, Application US/07838410
: Parent No. 5328635
: GENERAL INFORMATION:
: APPLICANT: SHIMA, NOBUYUKI
: APPLICANT: HIGASHIO, KANCI
: APPLICANT: NAGAO, MASAYA
: APPLICANT: COGAKI, FUMIKO
: APPLICANT: TAKAKURA, HIROAKI
: APPLICANT: TSUDA, EISUKE
: TITLE OF INVENTION: PLASMIDS CONTAINING DNA ENCODING THE
: TITLE OF INVENTION: AMINO ACID SEQUENCE OF TCF-II, TRANSFORMED CELLS WITH THE
: TITLE OF INVENTION: PLASMIDS, AND PRODUCTION METHOD OF BIOLOGICALLY ACTIVE
: TITLE OF INVENTION: SUBSTANCE USING THE TRANSFORMED CELLS
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MILLEN, WHITE, ZEILANO & BRANNIGAN, P.C.
: STREET: 2200 CLARENDON BLVD.
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22201
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```



```
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/838,410
/ FILING DATE: 19920311
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP91/00942
/ FILING DATE: 15-JUL-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 185852-1990
/ FILING DATE: 13-JUL-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: WHITE, JOHN L.
/ REGISTRATION NUMBER: 17,746
/ REFERENCE/DOCKET NUMBER: WAK 110
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-243-6333
/ TELEFAX: 703-243-6410
/ TELEX: 64191
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 723 amino acids
/ TYPE: AMINO ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ CELL TYPE: Fibroblast
/ FEATURE:
/ NAME/KEY: Domain
/ LOCATION: 393..405
/ OTHER INFORMATION: /note= "INTERNAL AMINO ACID
/ OTHER INFORMATION: SEQUENCE IN ALPHA-CHAIN"
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 490..505
/ OTHER INFORMATION: /note= "N-TERMINAL AMINO ACID
/ OTHER INFORMATION: SEQUENCE OF BETA-CHAIN"
/ FEATURE:
/ NAME/KEY: Domain
/ LOCATION: 605..623
/ OTHER INFORMATION: /note= "INTERNAL AMINO ACID
/ OTHER INFORMATION: SEQUENCE IN BETA-CHAIN"
US-07-838-410-1

Query Match          99.9%; Score 2576; DB 1; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTIHEFKSAKTTLLIKIDPALKITKRVNTADOCANRCTRNGGLPFTCKAFVPRK 60
DB 32 QRRGRRTIHEFKSAKTTLLIKIDPALKITKRVNTADOCANRCTRNGGLPFTCKAFVPRK 91
QY 61 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 120
DB 92 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 151
QY 121 WSGMIPIHEHSYRKDIOENYCRNPRGEGGPGCFSTNPVRYEVCDIPOCSEVECTCNG 180
DB 152 WSGMIPIHEHSYRKDIOENYCRNPRGEGGPGCFSTNPVRYEVCDIPOCSEVECTCNG 211
QY 181 ESRGLMDHTESGKICORMDHOTPHRHKFLPERYPDGKFPDNYCRNPDGQPRFWCTTLP 240
DB 212 ESRGLMDHTESGKICORMDHOTPHRHKFLPERYPDGKFPDNYCRNPDGQPRFWCTTLP 271
QY 241 HTRWEYCAIKTCADNTMNDTVPLEETTECIQGGEGYRGTVNTIANGIPQQRWDSQYPHE 300
DB 272 HTRWEYCAIKTCADNTMNDTVPLEETTECIQGGEGYRGTVNTIANGIPQQRWDSQYPHE 331
QY 301 HDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGCSDQIPNCDSHSGODCYTRNG 360
DB 332 HDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGCSDQIPNCDSHSGODCYTRNG 391
```

```
QY 361 KNYMGNLSQTRSGLTCSMWDKXNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGLTCSMWDKXNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPWCYTGN 451
QY 421 PLIPWDYCPISRCEGDTTPTIV 442
DB 452 PLIPWDYCPISRCEGDTTPTIV 473

RESULT 5
US-08-290-937B-1
/ Sequence 1, Application US/08290937B
/ Patent No. 5648233
/ GENERAL INFORMATION:
/ APPLICANT: YAMAGUCHI, KYOJI
/ APPLICANT: SHIMA, NOBUYUKI
/ APPLICANT: MURAKAMI, AKIHICO
/ APPLICANT: GOTO, MASAOKI
/ APPLICANT: TSUDA, EISUKE
/ APPLICANT: MASUNAGA, HIROAKI
/ APPLICANT: TAKAHIRA, REIKO
/ APPLICANT: OOGAKI, FUMIKO
/ APPLICANT: UEDA, MASATSUGU
/ APPLICANT: HIGASHIO, KANUJI
/ TITLE OF INVENTION: MODIFIED TCF
/ NUMBER OF SEQUENCES: 13
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Testa, Hurwitz & Thibault
/ STREET: 125 High St.
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/290,937B
/ FILING DATE: 19-AUG-1994
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CAMPBELL, PAULA A.
/ REGISTRATION NUMBER: 32,503
/ REFERENCE/DOCKET NUMBER: FJN-022
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 248-7000
/ TELEFAX: (617) 248-7100
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 723 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-290-937B-1

Query Match          99.9%; Score 2576; DB 1; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTIHEFKSAKTTLLIKIDPALKITKRVNTADOCANRCTRNGGLPFTCKAFVPRK 60
DB 32 QRRGRRTIHEFKSAKTTLLIKIDPALKITKRVNTADOCANRCTRNGGLPFTCKAFVPRK 91
QY 61 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 120
DB 92 ARKQCLMPFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKCOP 151
QY 121 WSGMIPIHEHSYRKDIOENYCRNPRGEGGPGCFSTNPVRYEVCDIPOCSEVECTCNG 180
```

Db 152 WSSMI PHEHSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCDDIPQCEVECMTCNG 211
QY 181 ESYRGLMDHTESGKICQWMDHOTPHRHKFLPERYPDKGFDNYYCRNPDGQPRMPCYTLDP 240
Db 212 ESYRGLMDHTESGKICQWMDHOTPHRHKFLPERYPDKGFDNYYCRNPDGQPRMPCYTLDP 271
QY 241 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGRGTNTIWNGLPCQRMDSQYPHE 300
Db 272 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGRGTNTIWNGLPCQRMDSQYPHE 331
QY 301 HMTPEPFCKDLRENYCRNPDGSESPMCFITTDNIRVGYCSQIPNCDMSHGQDCYRNG 360
Db 332 HMTPEPFCKDLRENYCRNPDGSESPMCFITTDNIRVGYCSQIPNCDMSHGQDCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGMPCYTCN 420
Db 392 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGMPCYTCN 451
QY 421 PLIPWDYCPISRCEGDTTPTIV 442
Db 452 PLIPWDYCPISRCEGDTTPTIV 473

RESULT 6

US-08-290-937B-2
Sequence 2, Application US/08290937B
Patent No. 5648233

GENERAL INFORMATION:

APPLICANT: YAMAGUCHI, KYOJI
APPLICANT: SHIMA, NOBUYUKI
APPLICANT: MORAKAMI, AKIHICO
APPLICANT: GOTO, MASAKI
APPLICANT: TSUDA, EISUKE
APPLICANT: MASUNAGA, HIROAKI
APPLICANT: TAKAHIRA, REIKO
APPLICANT: OOGAKI, FUMIKO
APPLICANT: UEDA, MASATSUGU
APPLICANT: HIGASHIO, KANUJI
TITLE OF INVENTION: MODIFIED TCF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,937B
FILING DATE: 19-AUG-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.
REGISTRATION/DOCKET NUMBER: 32,503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7100
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-937B-2

Query Match

99.9%; Score 2576; DB 1; Length 723;

Best Local Similarity 99.8%; Pred. No. 1,9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTTHEFKSKSKTLLIKIDPALKIKTKKNTVADOCANRCRNGLPFTCAFAVDK 60
Db 32 OKRRNTTHEFKSKSKTLLIKIDPALKIKTKKNTVADOCANRCRNGLPFTCAFAVDK 91
QY 61 ARKQCLWPFNMSGSGVKEFGHEFDLYENKDYIRNCIIGKRSYKTVSITKSGIKQP 120
Db 92 ARKQCLWPFNMSGSGVKEFGHEFDLYENKDYIRNCIIGKRSYKTVSITKSGIKQP 151
QY 121 WSSMI PHEHSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCDDIPQCEVECMTCNG 180
Db 152 WSSMI PHEHSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCDDIPQCEVECMTCNG 211
QY 181 ESYRGLMDHTESGKICQWMDHOTPHRHKFLPERYPDKGFDNYYCRNPDGQPRMPCYTLDP 240
Db 212 ESYRGLMDHTESGKICQWMDHOTPHRHKFLPERYPDKGFDNYYCRNPDGQPRMPCYTLDP 271
QY 241 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGRGTNTIWNGLPCQRMDSQYPHE 300
Db 272 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGRGTNTIWNGLPCQRMDSQYPHE 331
QY 301 HMTPEPFCKDLRENYCRNPDGSESPMCFITTDNIRVGYCSQIPNCDMSHGQDCYRNG 360
Db 332 HMTPEPFCKDLRENYCRNPDGSESPMCFITTDNIRVGYCSQIPNCDMSHGQDCYRNG 391
QY 361 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGMPCYTCN 420
Db 392 KNYMGNLSQTRSGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGMPCYTCN 451
QY 421 PLIPWDYCPISRCEGDTTPTIV 442
Db 452 PLIPWDYCPISRCEGDTTPTIV 473

RESULT 7

US-08-404-643-1
Sequence 1, Application US/08404643
Patent No. 5658742

GENERAL INFORMATION:

APPLICANT: HIGASHIO, KANUJI
APPLICANT: SHIMA, NOBUYUKI
APPLICANT: OOGAKI, FUMIKO
TITLE OF INVENTION: MONOCLONAL ANTIBODY
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: THIBEAULT
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,643
FILING DATE: 15-MAR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.
REGISTRATION/DOCKET NUMBER: 32,503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) -248-7000
TELEFAX: (617) -248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids

```

;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-404-643-1

```

Query Match	99.9%	Score 2576;	DB 1;	Length 723;
Best Local Similarity	99.8%	Pred. No. 1.9e-213;		
Matches 441; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	ERRKRRNTIHEFKFSATLTLLIKIDIPALKIKTKKVVNTADOCARCTRNKGLPTTCAAFYEDK	60
Dd	32 QQRGRNTIHEFKFSATLTLLIKIDIPALKIKTKKVVNTADOCARCTRNKGLPTTCAAFYEDK	91
Qy	61 ARKQCLMPFPNSMSGVCKKEFGHEFDLYENKDYINNCIIIGKRSYKTVSITTSKGIKOP	120
Dd	92 ARKQCLMPFPNSMSGVCKKEFGHEFDLYENKDYINNCIIIGKRSYKTVSITTSKGIKOP	151
Qy	121 WSMIMIPHEHSYRKDLOENYICRNPREEGEGPFCFINSNEVRYEYCDIPQSEYVCMTCNG	180
Dd	152 WSMIMIPHEHSYRKDLOENYICRNPREEGEGPFCFINSNEVRYEYCDIPQSEYVCMTCNG	211
Qy	181 ESVRGLMDHTESESKIQRMDHOTPHHKFLPERYPDKDFDNYNCNPDGQRPMPCTYLPD	240
Dd	212 ESVRGLMDHTESESKIQRMDHOTPHHKFLPERYPDKDFDNYNCNPDGQRPMPCTYLPD	271
Qy	241 HTRWEYCAIKTCADNTMDNDVPLETTECTIOGQSGEYGVNTIWMGIPCORMDSQYPHE	300
Dd	272 HTRWEYCAIKTCADNTMDNDVPLETTECTIOGQSGEYGVNTIWMGIPCORMDSQYPHE	331
Qy	301 HDMTPEPNFKCDLRENYICRNPDSSESPMCFITDPIIRIGYCSQIPNCMMSGQDCYVNG	360
Dd	332 HDMTPEPNFKCDLRENYICRNPDSSESPMCFITDPIIRIGYCSQIPNCMMSGQDCYVNG	391
Qy	361 KNTMGNLSQTRSGLTCSMMDKMDIEDLHRII FWEPDASKLLENYICRNPDDADHAGWCYTG	420
Dd	392 KNTMGNLSQTRSGLTCSMMDKMDIEDLHRII FWEPDASKLLENYICRNPDDADHAGWCYTG	451
Qy	421 PLIPMDYCPISRCBGDTTPYIV 442	
Dd	452 PLIPMDYCPISRCBGDTTPYIV 473	

```

RESULT 8
US-09-194-326-1
; Sequence 1, Application US/09194326
; Patent No. 6306827
GENERAL INFORMATION:
APPLICANT: Kinoshaki, Masahiko
APPLICANT: Ogawa, Hiroshi
APPLICANT: Masanaga, Hiroaki
APPLICANT: Kobayashi, Fumie
APPLICANT: Yamaguchi, Kyoji
APPLICANT: Higashio, Kenji
TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease
FILE REFERENCE: FJN-069
CURRENT APPLICATION NUMBER: US/09/194,326
CURRENT FILING DATE: 1998-11-24
EARLIER APPLICATION NUMBER: PCT/JP98/01221
EARLIER FILING DATE: 1998-03-20
EARLIER APPLICATION NUMBER: JP 94989
EARLIER FILING DATE: 1997-03-28
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 723
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: TCF-II
US-09-194-326-1

```

Best Local Similarity 99.8%; Pred. No. 1.9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	ERKBRNTHHEPKSKATLLIKITDPAALIKTKKVVNADOCANCTRNKGJLPTFCARFVDK	60
Db	32	QKRRNTHHEPKSKATLLIKITDPAALIKTKKVVNADOCANCTRNKGJLPTFCARFVDK	91
QY	61	ARKOCLWEPFMSMSGVAKKEFGHEFDLYENKDIYINCCIIIGKRSYGTATSIKSGIKOP	120
Db	92	ARKOCLWEPFMSMSGVAKKEFGHEFDLYENKDIYINCCIIIGKRSYGTATSIKSGIKOP	151
QY	121	WSSMIPHSHSTRGKLOENYCNPNPGEBCGPMCFIISNPEVRYEVCDIPOCSYEVCWTGNG	180
Db	152	WSSMIPHSHSTRGKLOENYCNPNPGEBCGPMCFIISNPEVRYEVCDIPOCSYEVCWTGNG	211
QY	181	ESYRGLMHTSSGKICQRMHDQTHPRHKLPERYPDKGFDNDYCRNPDOPRPMCTLLDP	240
Db	212	ESYRGLMHTSSGKICQRMHDQTHPRHKLPERYPDKGFDNDYCRNPDOPRPMCTLLDP	271
QY	241	HTRWEYCAIKTCADNTMMDTVPLETTEECIOQOGEGYRGTVNTIWNIGIPCORWDSQYPHE	300
Db	272	HTRWEYCAIKTCADNTMMDTVPLETTEECIOQOGEGYRGTVNTIWNIGIPCORWDSQYPHE	331
QY	301	HDMPENKCKDILRENYCRNPDGSSPMCFITDPNIRYGCQOIPNCDMSHODDCYRGNG	360
Db	332	HDMPENKCKDILRENYCRNPDGSSPMCFITDPNIRYGCQOIPNCDMSHODDCYRGNG	391
QY	361	KNYMGNISLQTRSSGLCSMMDKMKMEDLJHIIIEWEPASKLNEYCNRPDDDAHGPMWCTYGN	420
Db	392	KNYMGNISLQTRSSGLCSMMDKMKMEDLJHIIIEWEPASKLNEYCNRPDDDAHGPMWCTYGN	451
QY	421	PLIPWDCYPIRSCBGDTTPTIV 442	
Db	452	PLIPWDCYPIRSCBGDTTPTIV 473	

```

RESULT 9
US-09-600-991-18
; Sequence 18, Application US/09600991
; Patent No. 6551991
; GENERAL INFORMATION:
; APPLICANT: MEDICO, Enzo
; APPLICANT: MICHELI, Paolo
; APPLICANT: COLLESI, Chiara
; APPLICANT: CASELLI, Gianfranco
; APPLICANT: COMOGGIO, Paolo
; TITLE OF INVENTION: RECOMBINANT PROTEINS DERIVED FROM HGF AND MSP
; FILE REFERENCE: 0471-0162P
; CURRENT APPLICATION NUMBER: US/09/600,991
; CURRENT FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-600-991-18

```

Query Match	99.9%	Score 2576	DB 4	Length 723	
Best Local Similarity	99.8%	Pred. No. 1.9e+213			
Matches 441	Conservative 1	Mismatches 0	Indels 0	Gaps 0	
Qy	1	ERKRNTTHERKKSAKTLLIKIDPALKIKTKKVNTADOCANRCTNKGLPFTCKAFAVPDK	60		
	:				
Db	32	QRKRNTTHERKKSAKTLLIKIDPALKIKTKKVNTADOCANRCTNKGLPFTCKAFAVPDK	91		
Qy	61	ARKOCLWPEPMSMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKQCP	120		
	:				
Db	92	ARKOCLWPEPMSMSGGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVISITKSGIKQCP	151		
Qy	121	WSSMI PHEHSYRGKDLQENYICRNPRGEGGCPFTSNBRYEYVCDIPQCSYEVCCTCNG	180		
	:				
Db	152	WSSMI PHEHSYRGKDLQENYICRNPRGEGGCPFTSNBRYEYVCDIPQCSYEVCCTCNG	211		

```
OY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDDNYCRRNPDSQPRPWCYTLD 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDDNYCRRNPDSQPRPWCYTLD 271
OY 241 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNIGIPQCRWDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNIGIPQCRWDSQYPHE 331
OY 301 HMTPEENFKCDLRENYCRRNPDSQPRPWCFTTDPNIRVGYCSQIIPNCMSHGQDCYRGNG 360
DB 332 HMTPEENFKCDLRENYCRRNPDSQPRPWCFTTDPNIRVGYCSQIIPNCMSHGQDCYRGNG 391
OY 361 KNYMNLQSTRSGLTCSMWDKXMDLHRHIFWEPDASKLNNYCRNPDDAHGPMCTYGN 420
DB 392 KNYMNLQSTRSGLTCSMWDKXMDLHRHIFWEPDASKLNNYCRNPDDAHGPMCTYGN 451
OY 421 PLIPWDYCPISRCBGDTTPTIV 442
DB 452 PLIPWDYCPISRCBGDTTPTIV 473
```

```
RESULT 10
US-08-605-221-4
; Sequence 4, Application US/08605221
; Patent No. 6699837
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-B (five amino acids deletion type)
US-08-605-221-4
```

```
Query Match 99.9%; Score 2576; DB 4; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EKRKRNTHIEFKKSAKTTLIKIDPAKIKTKKVNTAQOCANRCTRNKGLPTCKAFVFDK 60
DB 32 QRRKRNTHIEFKKSAKTTLIKIDPAKIKTKKVNTAQOCANRCTRNKGLPTCKAFVFDK 91
OY 61 ARKQCLMPFPNMSGVYKKEFGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKCOP 120
DB 92 ARKQCLMPFPNMSGVYKKEFGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKCOP 151
OY 121 WSSMTIPHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVDIPQCEVEECMTNG 180
DB 152 WSSMTIPHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVDIPQCEVEECMTNG 211
OY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDDNYCRRNPDSQPRPWCYTLD 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDDNYCRRNPDSQPRPWCYTLD 271
OY 241 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNIGIPQCRWDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNIGIPQCRWDSQYPHE 331
OY 301 HMTPEENFKCDLRENYCRRNPDSQPRPWCFTTDPNIRVGYCSQIIPNCMSHGQDCYRGNG 360
DB 332 HMTPEENFKCDLRENYCRRNPDSQPRPWCFTTDPNIRVGYCSQIIPNCMSHGQDCYRGNG 391
OY 361 KNYMNLQSTRSGLTCSMWDKXMDLHRHIFWEPDASKLNNYCRNPDDAHGPMCTYGN 420
DB 392 KNYMNLQSTRSGLTCSMWDKXMDLHRHIFWEPDASKLNNYCRNPDDAHGPMCTYGN 451
```

```
OY 421 PLIPWDYCPISRCBGDTTPTIV 442
DB 452 PLIPWDYCPISRCBGDTTPTIV 473
```

```
RESULT 11
US-09-601-040A-10
; Sequence 10, Application US/09601040A
; Patent No. 6730657
; GENERAL INFORMATION:
; APPLICANT: MEDICO, Enzo et al.
; TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
; FILE REFERENCE: 0471-0161P
; CURRENT APPLICATION NUMBER: US/09/601,040A
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-601-040A-10
```

```
Query Match 99.9%; Score 2576; DB 4; Length 723;
Best Local Similarity 99.8%; Pred. No. 1.9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EKRKRNTHIEFKKSAKTTLIKIDPAKIKTKKVNTAQOCANRCTRNKGLPTCKAFVFDK 60
DB 32 QRRKRNTHIEFKKSAKTTLIKIDPAKIKTKKVNTAQOCANRCTRNKGLPTCKAFVFDK 91
OY 61 ARKQCLMPFPNMSGVYKKEFGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKCOP 120
DB 92 ARKQCLMPFPNMSGVYKKEFGHEFDLYENKDYIRNCIIGKRSYKGTVSITKSGIKCOP 151
OY 121 WSSMTIPHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVDIPQCEVEECMTNG 180
DB 152 WSSMTIPHEHSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVDIPQCEVEECMTNG 211
OY 181 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDDNYCRRNPDSQPRPWCYTLD 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIRHKFLPERYPDKGFDDNYCRRNPDSQPRPWCYTLD 271
OY 241 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNIGIPQCRWDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGQEGYRGVTNTIWNIGIPQCRWDSQYPHE 331
OY 301 HMTPEENFKCDLRENYCRRNPDSQPRPWCFTTDPNIRVGYCSQIIPNCMSHGQDCYRGNG 360
DB 332 HMTPEENFKCDLRENYCRRNPDSQPRPWCFTTDPNIRVGYCSQIIPNCMSHGQDCYRGNG 391
OY 361 KNYMNLQSTRSGLTCSMWDKXMDLHRHIFWEPDASKLNNYCRNPDDAHGPMCTYGN 420
DB 392 KNYMNLQSTRSGLTCSMWDKXMDLHRHIFWEPDASKLNNYCRNPDDAHGPMCTYGN 451
OY 421 PLIPWDYCPISRCBGDTTPTIV 442
DB 452 PLIPWDYCPISRCBGDTTPTIV 473
```

```
RESULT 12
US-09-601-040A-2
; Sequence 2, Application US/09601040A
; Patent No. 6730657
; GENERAL INFORMATION:
; APPLICANT: MEDICO, Enzo et al.
; TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
; FILE REFERENCE: 0471-0161P
; CURRENT APPLICATION NUMBER: US/09/601,040A
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
```

LENGTH: 729
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Recombinant sequence derived from Homo sapiens
US-09-601-040A-2

Query Match 99.9%; Score 2576; DB 4; Length 729;
Best Local Similarity 99.8%; Pred. No. 1.9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERRGRNTIHEFKSATTLLIKIDPALKIKTKKNTADQCANRCTRNGKGLPFTCKAFVFDK 60
DB 32 ORKRRNTIHEFKSATTLLIKIDPALKIKTKKNTADQCANRCTRNGKGLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSSSGVKKEFGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 120
DB 92 ARKQCLMPFPNSSSGVKKEFGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 151
QY 121 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDIPOCSEVECMTCNG 180
DB 152 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDIPOCSEVECMTCNG 211
QY 181 ESYRGLMDHTESGKICQRMWDHQTPIHRHKLPERYPDKGFDNNYCRNPDGAPRPWCYTLP 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIHRHKLPERYPDKGFDNNYCRNPDGAPRPWCYTLP 271
QY 241 HTRWEYCAIKTCADNTMNDTVPLETTETCIQGGEGYRGTVNTIWMGIPQQRWDSQYPHE 300
DB 272 HTRWEYCAIKTCADNTMNDTVPLETTETCIQGGEGYRGTVNTIWMGIPQQRWDSQYPHE 331
QY 301 HDWTPENFKCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDSHGODCYRGNG 360
DB 332 HDWTPENFKCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDSHGODCYRGNG 391
QY 361 KNYMGLSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGAPRPWCYTGN 420
DB 392 KNYMGLSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGAPRPWCYTGN 451
QY 421 PLIPWDYCPISRCGDTTPTIV 442
DB 452 PLIPWDYCPISRCGDTTPTIV 473

RESULT 13
US-09-601-040A-4
Sequence 4, Application US/09601040A
Patent No. 6730657
GENERAL INFORMATION:
APPLICANT: MEDICO, Enzo et al.
TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
FILE REFERENCE: 0471-0161P
CURRENT APPLICATION NUMBER: US/09/601, 040A
CURRENT FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 729
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Recombinant sequence derived from Homo sapiens
US-09-601-040A-4

Query Match 99.9%; Score 2576; DB 4; Length 729;
Best Local Similarity 99.8%; Pred. No. 1.9e-213;
Matches 441; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERRGRNTIHEFKSATTLLIKIDPALKIKTKKNTADQCANRCTRNGKGLPFTCKAFVFDK 60
DB 32 ORKRRNTIHEFKSATTLLIKIDPALKIKTKKNTADQCANRCTRNGKGLPFTCKAFVFDK 91
QY 61 ARKQCLMPFPNSSSGVKKEFGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 120

DB 92 ARKQCLMPFPNSSSGVKKEFGHEFDLYENKDIYRNCIIIGKRSYKGTVSITKSGIKCOP 151
QY 121 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDIPOCSEVECMTCNG 180
DB 152 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDIPOCSEVECMTCNG 211
QY 181 ESYRGLMDHTESGKICQRMWDHQTPIHRHKLPERYPDKGFDNNYCRNPDGAPRPWCYTLP 240
DB 212 ESYRGLMDHTESGKICQRMWDHQTPIHRHKLPERYPDKGFDNNYCRNPDGAPRPWCYTLP 271
QY 241 HTRWEYCAIKTCADNTMNDTVPLETTETCIQGGEGYRGTVNTIWMGIPQQRWDSQYPHE 300
DB 272 HTRWEYCAIKTCADNTMNDTVPLETTETCIQGGEGYRGTVNTIWMGIPQQRWDSQYPHE 331
QY 301 HDWTPENFKCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDSHGODCYRGNG 360
DB 332 HDWTPENFKCKDLRENYCRNPDGSESPWCFSTDPNIRVGYCSQIPNCDSHGODCYRGNG 391
QY 361 KNYMGLSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGAPRPWCYTGN 420
DB 392 KNYMGLSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGAPRPWCYTGN 451
QY 421 PLIPWDYCPISRCGDTTPTIV 442
DB 452 PLIPWDYCPISRCGDTTPTIV 473

RESULT 14
US-08-030-410-3
Sequence 3, Application US/08030410
Patent No. 6221359
GENERAL INFORMATION:
APPLICANT: Komiya, Atsushi
APPLICANT: Nakahata, Tatsuroshi
APPLICANT: Kubo, Teisuo
APPLICANT: Tanaka, Ryuhai
APPLICANT: Kawano, Genji
APPLICANT: Sudo, Teisuo
APPLICANT: Sano, Emiko
APPLICANT: Kojima, Katsunaki
TITLE OF INVENTION: HEMOPOIETIC STEM CELL MULTIPLIER
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/030,410
FILING DATE: 19930521
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8898
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-030-410-3

Query Match 99.8%; Score 2573; DB 3; Length 723;
 Best Local Similarity 99.5%; Pred. No. 3 4e-213;
 Matches 440; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKITKKVNTADOCANRCTRNGKLPFTCKAFVFDK 60
DB 32 QKRKRNTIHEFKSAKTTLIKIDPALKITKKVNTADOCANRCTRNGKLPFTCKAFVFDK 91
QY 61 ARKQCLMPFNSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKOP 120
DB 92 ARKQCLMPFNSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKOP 151
QY 121 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVREYVCDIPQCSVEECMTGNG 180
DB 152 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVREYVCDIPQCSVEECMTGNG 211
QY 161 ESYRGLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 271
QY 241 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGGGSGYGTNTIWNGLPCQRMDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGGGSGYGTNTIWNGLPCQRMDSQYPHE 331
QY 301 HDMTEPNFKCDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHSGQDCYRNG 360
DB 332 HDMTEPNFKCDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHSGQDCYRNG 391
QY 361 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGPWCYTGN 451
QY 421 PLIPWDYCPISRCEGDTPTIV 442
DB 452 PLIPWDYCPISRCEGDTPTIV 473

```

RESULT 15

US-08-290-937B-3

Sequence 3, Application US/08230937B

Patent No. 5648233

GENERAL INFORMATION:

APPLICANT: YAMAGUCHI, KYOJI

APPLICANT: SHIMA, NOBUYUKI

APPLICANT: MURAKAMI, AKIHICO

APPLICANT: GOTO, MASAKI

APPLICANT: TSUDA, EISUKE

APPLICANT: MASUNAGA, HIROAKI

APPLICANT: TAKAHIRA, REIKO

APPLICANT: OOGAKI, FUMIKO

APPLICANT: UEDA, MASATOSUGU

APPLICANT: HIGASHIO, KANJI

TITLE OF INVENTION: MODIFIED TCF

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Testa, Hurwitz & Thibault

STREET: 125 High St.

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,937B

FILING DATE: 19-AUG-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: CAMPBELL, PAULA A.
 REGISTRATION NUMBER: 32,503
 REFERENCE/DOCKET NUMBER: FJN-022
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 723 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-290-937B-3

Query Match 99.7%; Score 2570; DB 1; Length 723;
 Best Local Similarity 99.5%; Pred. No. 6 1e-213;
 Matches 440; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKITKKVNTADOCANRCTRNGKLPFTCKAFVFDK 60
DB 32 QKRKRNTIHEFKSAKTTLIKIDPALKITKKVNTADOCANRCTRNGKLPFTCKAFVFDK 91
QY 61 ARKQCLMPFNSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKOP 120
DB 92 ARKQCLMPFNSMSGVKKFGEHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKOP 151
QY 121 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVREYVCDIPQCSVEECMTGNG 180
DB 152 WSSMIPHEHSYRGKDLQENYCRNPRGEGGPMCFSTNPEVREYVCDIPQCSVEECMTGNG 211
QY 161 ESYRGLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 240
DB 212 ESYRGLMDHTESGKICQRMWHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDP 271
QY 241 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGGGSGYGTNTIWNGLPCQRMDSQYPHE 300
DB 272 HTRMEYCAIKTCADNTMNDTVPLETTTCIOGGGSGYGTNTIWNGLPCQRMDSQYPHE 331
QY 301 HDMTEPNFKCDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHSGQDCYRNG 360
DB 332 HDMTEPNFKCDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIPNCDSHSGQDCYRNG 391
QY 361 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGPWCYTGN 420
DB 392 KNYMGNLSQTRSGGLTCSMWDKXMEDLHRHIFWEPDASKLNNYCRNPDGDAHGPWCYTGN 451
QY 421 PLIPWDYCPISRCEGDTPTIV 442
DB 452 PLIPWDYCPISRCEGDTPTIV 473

```

Search completed: October 3, 2005, 06:53:49
 Job time : 1361.31 secs